

Result No.	Score	Query Match	Length	DB ID	Description
1	38	46.9	20	1	R92907 HLA-B*2702 CTL modu
2	38	46.9	20	1	R95428 HLA-B*2702 84-75-84
3	38	46.9	20	1	W33778 Immunomodulating d
4	38	46.9	513	1	R79945 Helicobacter pylori
5	36	44.4	176	1	W55346 H. pylori ORF 07ep
6	36	44.4	178	1	Y11022 H. pylori ORF 02ge
7	36	44.4	852	1	W9065 FIV-NC8UL clone JS
8	36	44.4	854	1	R44957 Feline Immunodefici
9	36	44.4	854	1	R51250 FIV PR envelope p
10	36	44.4	855	1	R51255 FIV SWISS22 envelo
11	36	44.4	855	1	R59585 FIV UK8 envelope p
12	36	44.4	856	1	R51247 FIV envelope prote
13	36	44.4	856	1	R51248 FIV petalum envelop
14	36	44.4	856	1	R51249 FIV PRT-Fl4 envelo
15	36	44.4	856	1	R51254 FIV DUTCH19K1 enve
16	36	44.4	856	1	W53671 FIV PR clone 34 O
17	36	44.4	870	1	R51246 Composite sequence
18	36	44.4	1287	1	R79944 Helicobacter pylori
19	36	44.4	1288	1	W55547 H. pylori ORF 14ee
20	36	44.4	1288	1	W56685 H. pylori ORF 07ee
21	36	44.4	1296	1	R41198 CT. Recombinant He
22	35	43.2	20	1	R95430 HLA-B*2702 84-75T/7
23	35	43.2	207	1	R32202 Broad bean leaf PP
24	35	43.2	267	1	W32427 Mycobacterium tube
25	35	43.2	267	1	W32359 Mycobacterium tube
26	35	43.2	267	1	W64299 Mycobacterium tube
27	35	43.2	267	1	W81662 M. tuberculosis im
28	35	43.2	435	1	R32999 Rat choline kinase
29	35	43.2	511	1	P81161 Recombinant alpha-
30	35	43.2	511	1	P81180 Sequence of alpha-
31	35	43.2	512	1	R07574 Alpha-amylase enco
32	35	43.2	855	1	R28033 Env polypeptide of
33	35	43.2	855	1	R51251 FIV JAPAN7M2 envel
34	35	43.2	855	1	R51252 FIV UK2 envelope p





PR 02-APR-1996; US-758731.  
 PR 25-OCT-1996; US-736905.  
 PR 28-OCT-1996; US-738859.  
 PA (ASTR.) ASTRA AB.  
 PI Alm RA, Smith D;  
 DR WPI; 97-503122/46.  
 DR N-PSDB; V24755.

PT Helicobacter pylori nucleic acid sequences and encoded  
 PT polypeptide(s) - useful in vaccines to treat or prevent H. pylori  
 PT infection and for diagnosis of H. pylori infection  
 PS Claim 14; Page 570-571; 1145pp; English.  
 CC This sequence is a H. pylori protein of unspecified function.  
 CC The protein may be used in a vaccine to prevent or treat H. pylori  
 CC infection or to identify H. pylori polypeptide binding compounds,  
 CC useful as potential H. pylori life cycle activators or inhibitors. The  
 CC DNA and probes derived from it may be used for the identification of  
 CC H. pylori in a sample and the diagnosis of H. pylori infection. Nucleic  
 CC acid sequences complementary to the DNA act as antisense sequences and  
 CC can be used to prevent the translation of H. pylori mRNA. Antibodies  
 CC against the protein can be used in immunoassays to evaluate the abundance  
 CC and distribution of H. pylori-specific antigens. The genomic sequence of  
 CC H. pylori (ATCC 55679) was determined from overlapping contigs generated  
 CC by mechanically shearing the bacterial DNA. The sequences were analysed  
 CC for ORF of at least 180 nucleotides, and the predicted coding regions  
 CC defined by computer evaluation. To identify likely H. pylori antigens for  
 CC vaccine development, the amino acid sequences predicted from various ORF  
 CC were analysed for significant homology to other known or exported  
 CC membrane proteins. Having identified and determined the sequences of  
 CC interest, particular regions can be isolated from H. pylori by PCR  
 CC amplification for recombinant polypeptide production, e.g. in E. coli  
 CC hosts.  
 SQ Sequence 176 AA;

Query Match 44.4%; Score 36; DB 1; Length 176;  
 Best Local Similarity 46.7%; Pred. No. 40;  
 Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 2 YKLVKINNIRIVK 16  
 || :|||: : :||  
 DB 16 YKPLKINSADLLK 30

## RESULT 6

ID Y11022 standard; Protein; 178 AA.

AC Y11022; 1999 (first entry).  
 DE H. pylori ORF 02gei0116\_15632000\_c2\_114 cell envelope protein.  
 KW Vaccine; probe; diagnostic; ORF; cell envelope protein;  
 KW secreted protein; cytoplasmic protein; cellular protein.  
 OS Helicobacter pylori.  
 PN W09824475-AL.

PD 11-JUN-1998.  
 PF 05-DEC-1997; U22104.  
 PR 14-JUL-1997; US-891928.  
 PR 05-DEC-1996; US-759625.  
 PR 25-MAR-1997; US-823745.

PA (ASTR.) ASTRA AB.  
 PI Alm RA, Castriotta LM, Doig PC, Kabok Z, Smith D;  
 DR WPI; 98-333051/29.  
 DR N-PSDB; X30551.

PT New isolated Helicobacter pylori nucleic acids - used to develop  
 PT products for the diagnosis, prevention and treatment of infection by  
 PT H. pylori and other Helicobacter species  
 PS Claims 37, 41; Page 184; 339pp; English.  
 CC Recombinant or substantially pure preparations of H. pylori polypeptides  
 CC are disclosed, together with the nucleic acids encoding them. In all,  
 CC 97 ORFs are shown. The proteins are variously cell envelope proteins,  
 CC cytoplasmic proteins, secreted proteins or other cellular proteins.  
 CC Vaccines containing the nucleic acids or proteins are claimed, as are  
 CC probes containing at least 8 nucleotides from the nucleic acid  
 CC sequences. The vaccines are useful for treating or reducing the risk of

CC H. pylori infections, and the probes can be used diagnostically for  
 CC detecting the presence of Helicobacter in a sample. The products are  
 CC also of use in screening for compounds having the ability to interfere  
 CC with the H. pylori life cycle or to inhibit H. pylori infection.  
 SQ Sequence 178 AA;

Query Match 44.4%; Score 36; DB 1; Length 178;  
 Best Local Similarity 50.0%; Pred. No. 41;  
 Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 3 KLVKINNIRIVK 16  
 | | :||| : |  
 DB 90 KSVLPNNIEIKK 103

## RESULT 7

W79065  
 ID W79065 standard; Protein; 852 AA.

AC W79065;  
 DT 11-JAN-1999 (first entry)  
 DE FIV-NCSU1 clone JSY3 envelope protein (env).  
 KW FIV; retrovirus; immunodeficiency; diagnosis; vaccine;  
 KW envelope protein; env; cat; animal model.  
 OS Feline immunodeficiency virus isolate NCSU1 clone JSY3.  
 FH Key Location/Qualifiers  
 FT Domain 691..702  
 FT /label= TM  
 FT /note= "transmembrane domain (Claim 19)"

PN W09839451-AL.

PD 11-SEP-1998.  
 PF 04-MAR-1998; U04147.  
 PR 05-MAR-1997; US-811682.  
 PA (UYNC-) UNIV NORTH CAROLINA STATE.  
 PI Tompkins WB, Tompkins WAF, Yang J;  
 DR WPI; 98-506365/43.  
 DR N-PSDB; V57355-56.  
 PT Feline immunodeficiency virus clone, JSY3 - retaining in vivo  
 PT characteristics of parent virus, useful e.g. to produce vaccines to  
 PT protect felines against feline immunodeficiency virus  
 PS Claim 19; Page 44-47; 128pp; English.  
 CC This is the amino acid sequence of the FIV isolate NCSU1 clone JSY3  
 CC envelope protein (env). The sequence was deduced from JSY3  
 CC proviral DNA (see V57355-56). The JSY3 clone, polypeptides and  
 CC encoding nucleic acids are useful for producing vaccines to protect  
 CC feline species against FIV; such vaccines may contain e.g. live  
 CC attenuated JSY3 virus, antigenic fragments (e.g. the env  
 CC transmembrane peptide), or host cells transformed with nucleic  
 CC acids and expressing virus antigen on the surface etc. They are  
 CC also useful diagnostically. Since JSY3 retains the essential in  
 CC vivo biological characteristics of the parent virus (including  
 CC induction of immunodeficiency and inversion of the CD4+CD8+ ratio  
 CC by 6 weeks post infection), it is useful for producing animal  
 CC models for studying the virus and other retroviral diseases, e.g.  
 CC HIV-1 infection, and testing possible treatments.  
 SQ Sequence 852 AA;

Query Match 44.4%; Score 36; DB 1; Length 852;  
 Best Local Similarity 66.7%; Pred. No. 2.2e+02;  
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 6 IKINNIRIV 14  
 :|||:|:|  
 DB 656 LKINNRLRV 664

## RESULT 8

R44957  
 ID R44957 standard; Protein; 854 AA.

AC R44957;  
 DT 05-JUL-1994 (first entry)  
 DE Feline Immunodeficiency Virus (Wo isolate) ENV protein.

KW FIV; French virus strain; Wo isolate; envelope glycoprotein;  
 KW feline acquired immunodeficiency syndrome; cat; vaccine;  
 OS Retrovirus; lentivirus.  
 FH Feline immunodeficiency virus (Wo isolate).  
 FT Key Location/Qualifiers  
 FT 26..72  
 FT /label= V1  
 FT /note= "variable region capable of inducing a  
 FT Wo strain-specific immune response"  
 FT 96..174  
 FT /label= V2  
 FT /note= "variable region capable of inducing a  
 FT Wo strain-specific immune response"  
 FT 253..289  
 FT /label= SU1  
 FT /note= "part of external surface glycoprotein"  
 FT 388..424  
 FT /label= SU2  
 FT /note= "part of external surface glycoprotein"  
 FT 361..422  
 FT /label= V3  
 FT /note= "variable region capable of inducing a  
 FT Wo strain-specific immune response"  
 FT 467..492  
 FT /label= SU3  
 FT /note= "part of external surface glycoprotein"  
 FT 452..487  
 FT /label= V4  
 FT /note= "variable region capable of inducing a  
 FT Wo strain-specific immune response"  
 FT 508..528  
 FT /label= SU4  
 FT /note= "part of external surface glycoprotein"  
 FT 541..570  
 FT /label= V5  
 FT /note= "variable region capable of inducing a  
 FT Wo strain-specific immune response"  
 FT 572..606  
 FT /label= SU5  
 FT /note= "part of external surface glycoprotein"  
 FT 586..612  
 FT /label= V6  
 FT /note= "variable region capable of inducing a  
 FT Wo strain-specific immune response"  
 FT 595..647  
 FT /label= TM1  
 FT /note= "part of transmembrane glycoprotein"  
 FT 681..711  
 FT /label= TM2  
 FT /note= "part of transmembrane glycoprotein"  
 FT 710..718  
 FT /label= V7  
 FT /note= "variable region capable of inducing a  
 FT Wo strain-specific immune response"  
 FT 765..778  
 FT /label= V8  
 FT /note= "variable region capable of inducing a  
 FT Wo strain-specific immune response"  
 FT 832..842  
 FT /label= V9  
 FT /note= "variable region capable of inducing a  
 FT Wo strain-specific immune response"  
 EP-577458-A.  
 PN 05-JAN-1994. 401538.  
 PD 16-JUN-1993; FR-007257.  
 PR 16-JUN-1992; FR-007257.  
 PR 23-NOV-1992; FR-014036.  
 PA (CNRS ) CNRS CENT NAT RECH SCI.  
 PI Avrameas A, Chappay C, Hurtrel B, Klatzmann D, Moraillon A;  
 PI Pincino G, Saurin W, Sonigo P, Strosberg AD;  
 DR WPI: 94-005689/02.  
 DR N-PSDB; Q54396.

PT Nucleic acid sequences of GAG and ENV genes - of feline immune  
 PT deficiency virus, strain Wo, and derived probes, primers, vectors  
 PT and peptide(s), for diagnosis and vaccination  
 PS Claim 20; Page 34-36; 56pp; French.  
 CC Peripheral blood mononuclear cells (PBMC) from cats naturally  
 CC infected with the FIV French strain Wo were co-cultured with PBMC  
 CC from uninfected cats in the presence of recombinant human IL-2 and  
 CC concanavalin A. After 15 days, total cellular DNA was extracted  
 CC from cells positive for Mg-dependent reverse transcriptase activity  
 CC and was used as starting material for amplifying FIV env and gag  
 CC genes. PCR primers were designed based on conserved regions in the  
 CC Petalum isolate of FIV. The env gene was obtained from 3  
 CC overlapping subfragments. Sequence analysis and comparison with  
 CC other FIV isolates identified 9 variable regions suitable for  
 CC designing strain-specific diagnostic and therapeutic agents.  
 SQ Sequence 854 AA;  
 QY 6 IKINNIRIV 14  
 DB 658 LKINNRLV 666  
 Query Match 44.4%; Score 36; DB 1; Length 854;  
 Best Local Similarity 66.7%; Pred. NO. 2.2e+02;  
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
 RESULT 9  
 R51250  
 ID R51250 standard; peptide; 854 AA.  
 AC R51250;  
 DT 08-OCT-1994 (first entry)  
 DE FIV PPR envelope protein sequence.  
 KW Feline immunodeficiency virus; FIV; vaccine; diagnostic; AIDS;  
 KW T-lymphotropic lentivirus; FIV UR8; FIV Petalum; envelope protein.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT peptide 1..178  
 FT region /note= "Hydrophobic leader peptide"  
 FT 51..66  
 FT /note= "V1 region"  
 FT 95..173  
 FT /note= "V2 region"  
 FT 179..608  
 FT /note= "Surface glycoprotein"  
 FT 360..424  
 FT /note= "V3 region"  
 FT 451..483  
 FT /note= "V4 region"  
 FT 539..565  
 FT /note= "V5 region"  
 FT 609..854  
 FT /note= "Transmembrane protein"  
 PN W09406471-A.  
 PD 31-MAR-1994.  
 PE 20-SEP-1993; G01974.  
 PR 21-SEP-1992; GB-019936.  
 PA (PITM ) PITMAN MOORE INC.  
 PI Francis MJ;  
 DR WPI: 94-118168/14.  
 PT Feline immunodeficiency virus antigenic polypeptide(s) and  
 PT nucleic acid - used to prepare prods. for combating or diagnosis  
 PT of feline immunodeficiency infection  
 PS Disclosure; Figure 2; 68pp; English.  
 CC The sequences given in R51248-R51262 and R58584-85 represent the  
 CC feline immunodeficiency virus (FIV) envelope protein from different  
 CC strains and serotypes. The consensus sequence based on these, is given  
 CC in R51247. The synthetic FIV peptides of the invention were derived  
 CC principally from a combination of the sequences of the FIV UK 8 and  
 CC Petalum isolates, a composite sequence of which is given in R51246.  
 CC Variations in the sequence may occur between different strains or  
 CC serotypes, isolates of different geographical origin or even between  
 CC different isolates from the same host.

SQ Sequence 854 AA;

Query Match 44.4%; Score 36; DB 1; Length 854;

Best Local Similarity 66.7%; Pred. No. 2.2e+02;

Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 6 IKINNIRIV 14

:||||:|

Db 658 LKINNRLV 666

RESULT 10

R51255

ID R51255 standard; peptide; 855 AA.

AC R51255;

DT 08-OCT-1994 (first entry)

DE FIV SWISS22 envelope protein sequence.

KW Feline immunodeficiency virus; FIV; vaccine; diagnostic; AIDS;

KW T-lymphotropic lentivirus; FIV UK8; FIV Petaluma; envelope protein.

OS Synthetic.

FH Key Location/Qualifiers

FT peptide 1..178

FT /note= "Hydrophobic leader peptide"

FT region 51..66

FT /note= "v1 region"

FT region 95..173

FT /note= "v2 region"

FT protein 179..609

FT /note= "Surface glycoprotein"

FT region 360..424

FT /note= "v3 region"

FT region 451..483

FT /note= "v4 region"

FT region 539..566

FT /note= "v5 region"

FT protein 610..855

FT /note= "Transmembrane protein"

FT WO9406471-A.

PN 31-MAR-1994.

PD 20-SEP-1993; G01974.

PR 21-SEP-1992; GB-019936.

PA (PTM ) PITMAN MOORE INC.

PI Francis MJ;

DR WPI; 94-118168/14.

PT Feline immunodeficiency virus antigenic polypeptide(s) and

PT nucleic acid - used to prepare prods. for combating or diagnosis

PT of feline immunodeficiency infection

PS Disclosure; Figure 2; 68pp; English.

CC The sequences given in R51248-R51262 and R58584-85 represent the

CC feline immunodeficiency virus (FIV) envelope protein from different

CC strains and serotypes. The consensus sequence based on these, is given

CC in R51247. The synthetic FIV peptides of the invention were derived

CC principally from a combination of the sequences of the FIV UK 8 and

CC Petaluma isolates, a composite sequence of which is given in R51246.

CC Variations in the sequence may occur between different strains or

CC serotypes, isolates of different geographical origin or even between

CC different isolates from the same host.

SQ Sequence 855 AA;

Query Match

Best Local Similarity 44.4%; Score 36; DB 1; Length 855;

Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 6 IKINNIRIV 14

:||||:|

Db 659 LKINNRLV 667

RESULT 11

R58585

ID R58585 standard; peptide; 855 AA.

AC

DT 08-OCT-1994 (first entry)

DE FIV UK8 envelope protein sequence.

KW Feline immunodeficiency virus; FIV; vaccine; diagnostic; AIDS;

KW T-lymphotropic lentivirus; FIV UK8; FIV Petaluma; envelope protein.

OS Synthetic.

FH Key Location/Qualifiers

FT peptide 1..178

FT /note= "Hydrophobic leader peptide"

FT region 51..66

FT /note= "v1 region"

FT region 95..173

FT /note= "v2 region"

FT protein 179..609

FT /note= "Surface glycoprotein"

FT region 360..424

FT /note= "v3 region"

FT region 451..483

FT /note= "v4 region"

FT region 539..566

FT /note= "v5 region"

FT protein 610..855

FT /note= "Transmembrane protein"

FT WO9406471-A.

PN 31-MAR-1994.

PD 20-SEP-1993; G01974.

PR 21-SEP-1992; GB-019936.

PA (PTM ) PITMAN MOORE INC.

PI Francis MJ;

DR WPI; 94-118168/14.

PT Feline immunodeficiency virus antigenic polypeptide(s) and

PT nucleic acid - used to prepare prods. for combating or diagnosis

PT of feline immunodeficiency infection

PS Disclosure; Figure 2; 68pp; English.

CC The sequences given in R51248-R51262 and R58584-85 represent the

CC feline immunodeficiency virus (FIV) envelope protein from different

CC strains and serotypes. The consensus sequence based on these, is given

CC in R51247. The synthetic FIV peptides of the invention were derived

CC principally from a combination of the sequences of the FIV UK 8 and

CC Petaluma isolates, a composite sequence of which is given in R51246.

CC Variations in the sequence may occur between different strains or

CC serotypes, isolates of different geographical origin or even between

CC different isolates from the same host.

SQ Sequence 855 AA;

Query Match

Best Local Similarity 44.4%; Score 36; DB 1; Length 855;

Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 6 IKINNIRIV 14

:||||:|

Db 659 LKINNRLV 667

RESULT 12

R51247

ID R51247 standard; peptide; 856 AA.

AC R51247;

DT 08-OCT-1994 (first entry)

DE FIV envelope protein consensus sequence.

KW Feline immunodeficiency virus; FIV; vaccine; diagnostic; AIDS;

KW T-lymphotropic lentivirus; FIV UK8; FIV Petaluma; envelope protein.

OS Synthetic.

FH Key Location/Qualifiers

FT peptide 1..178

FT /note= "Hydrophobic leader peptide"

FT region 51..66

FT /note= "v1 region"

FT region 95..173

FT /note= "v2 region"

FT protein 179..610

FT /note= "Surface glycoprotein"

FT region 360..424  
 FT /note= "V3 region"  
 FT 451..483  
 FT /note= "V4 region"  
 FT 539..567  
 FT /note= "V5 region"  
 FT 611..856  
 FT /note= "Transmembrane protein"

PN WO9406471-A.  
 PD 31-MAR-1994.  
 PD 20-SEP-1993; G01974.  
 PR 21-SEP-1992; GB-019936.  
 PA (PITM ) PITMAN MOORE INC.  
 PI Francis MJ;  
 DR WPI; 94-118168/14.

PT Feline immunodeficiency virus antigenic polypeptide(s) and  
 PT nucleic acid - used to prepare prods. for combating or diagnosis  
 PT of feline immunodeficiency infection  
 PS Disclosure; Figure 2; 68pp; English.  
 CC The sequences given in R51248-R51262 and R58584-85 represent the  
 CC feline immunodeficiency virus (FIV) envelope protein from different  
 CC strains and serotypes. The consensus sequence based on these, is given  
 CC in R51247. The synthetic FIV peptides of the invention were derived  
 CC principally from a combination of the sequences of the FIV UK 8 and  
 CC Petaluma isolates, a composite sequence of which is given in R51246.  
 CC Variations in the sequence may occur between different strains or  
 CC serotypes, isolates of different geographical origin or even between  
 CC different isolates from the same host.  
 CC Sequence 856 AA;  
 SQ

Query Match 44.4%; Score 36; DB 1; Length 856;  
 Best Local Similarity 66.7%; Pred. No. 2.2e+02;  
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 6 IKINNIRIV 14  
 :||||:|  
 Db 660 LKINNLRLV 668

## RESULT 13

ID R51248 standard; peptide; 856 AA.  
 AC R51248;  
 DT 08-OCT-1994 (first entry)  
 DE FIV petaluma envelope protein sequence.  
 KW Feline immunodeficiency virus; FIV; vaccine; diagnostic; AIDS;  
 KW T-lymphotropic lentivirus; FIV UK8; FIV Petaluma; envelope protein.  
 OS Synthetic.

FT Key Location/Qualifiers  
 FT peptide 1..178  
 FT /note= "Hydrophobic leader peptide"  
 FT region 51..66  
 FT /note= "V1 region"  
 FT region 95..173  
 FT /note= "V2 region"  
 FT protein 179..610  
 FT /note= "Surface glycoprotein"  
 FT region 360..424  
 FT /note= "V3 region"  
 FT region 451..483  
 FT /note= "V4 region"  
 FT region 539..567  
 FT /note= "V5 region"  
 FT protein 611..856  
 FT /note= "Transmembrane protein"

PN WO9406471-A.  
 PD 31-MAR-1994.  
 PD 20-SEP-1993; G01974.  
 PR 21-SEP-1992; GB-019936.  
 PA (PITM ) PITMAN MOORE INC.  
 PI Francis MJ;  
 DR WPI; 94-118168/14.

PT Feline immunodeficiency virus antigenic polypeptide(s) and  
 PT nucleic acid - used to prepare prods. for combating or diagnosis  
 PT of feline immunodeficiency infection  
 PS Disclosure; Figure 2; 68pp; English.  
 CC The sequences given in R51248-R51262 and R58584-85 represent the  
 CC feline immunodeficiency virus (FIV) envelope protein from different  
 CC strains and serotypes. The consensus sequence based on these, is given  
 CC in R51247. The synthetic FIV peptides of the invention were derived  
 CC principally from a combination of the sequences of the FIV UK 8 and  
 CC Petaluma isolates, a composite sequence of which is given in R51246.  
 CC Variations in the sequence may occur between different strains or  
 CC serotypes, isolates of different geographical origin or even between  
 CC different isolates from the same host.  
 CC Sequence 856 AA;  
 SQ

Query Match 44.4%; Score 36; DB 1; Length 856;  
 Best Local Similarity 66.7%; Pred. No. 2.2e+02;  
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 6 IKINNIRIV 14  
 :||||:|  
 Db 660 LKINNLRLV 668

## RESULT 14

ID R51249 standard; peptide; 856 AA.  
 AC R51249;  
 DT 08-OCT-1994 (first entry)  
 DE FIV PET-F14 envelope protein sequence.  
 KW Feline immunodeficiency virus; FIV; vaccine; diagnostic; AIDS;  
 KW T-lymphotropic lentivirus; FIV UK8; FIV Petaluma; envelope protein.  
 OS Synthetic.

FT Key Location/Qualifiers  
 FT peptide 1..178  
 FT /note= "Hydrophobic leader peptide"  
 FT region 51..66  
 FT /note= "V1 region"  
 FT region 95..173  
 FT /note= "V2 region"  
 FT protein 179..610  
 FT /note= "Surface glycoprotein"  
 FT region 360..424  
 FT /note= "V3 region"  
 FT region 451..483  
 FT /note= "V4 region"  
 FT region 539..567  
 FT /note= "V5 region"  
 FT protein 611..856  
 FT /note= "Transmembrane protein"

PN WO9406471-A.  
 PD 31-MAR-1994.  
 PD 20-SEP-1993; G01974.  
 PR 21-SEP-1992; GB-019936.  
 PA (PITM ) PITMAN MOORE INC.  
 PI Francis MJ;  
 DR WPI; 94-118168/14.

PT Feline immunodeficiency virus antigenic polypeptide(s) and  
 PT nucleic acid - used to prepare prods. for combating or diagnosis  
 PT of feline immunodeficiency infection  
 PS Disclosure; Figure 2; 68pp; English.  
 CC The sequences given in R51248-R51262 and R58584-85 represent the  
 CC feline immunodeficiency virus (FIV) envelope protein from different  
 CC strains and serotypes. The consensus sequence based on these, is given  
 CC in R51247. The synthetic FIV peptides of the invention were derived  
 CC principally from a combination of the sequences of the FIV UK 8 and  
 CC Petaluma isolates, a composite sequence of which is given in R51246.  
 CC Variations in the sequence may occur between different strains or  
 CC serotypes, isolates of different geographical origin or even between  
 CC different isolates from the same host.  
 CC Sequence 856 AA;  
 SQ

Query Match 44.4%; Score 36; DB 1; Length 856;  
 Best Local Similarity 66.7%; Pred. NO. 2.2e+02;  
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 6 IKINNIRIV 14  
 :||||:|  
 Db .660 LKINLRV 668

## RESULT 15

R51254  
 ID R51254 standard; peptide: 856 AA.  
 AC R51254;  
 DE 08-OCT-1994 (first entry)  
 DE FIV DUTCH19K1 envelope protein sequence.  
 KW Feline immunodeficiency virus; FIV; vaccine; diagnostic; AIDS;  
 KW T-lymphotropic lentivirus; FIV UK8; FIV Petaluma; envelope protein.  
 OS Synthetic.  
 FH Key  
 FT peptide  
 FT region /note= "Hydrophobic leader peptide"  
 FT region 51..66  
 FT region /note= "v1 region"  
 FT region 95..173  
 FT region /note= "v2 region"  
 FT protein 179..610  
 FT region /note= "Surface glycoprotein"  
 FT region 360..424  
 FT region /note= "v3 region"  
 FT region 451..483  
 FT region /note= "v4 region"  
 FT region 539..567  
 FT protein /note= "v5 region"  
 FT protein 611..856  
 FT /note= "Transmembrane protein"  
 PN W09406471-A.  
 PD 31-MAR-1994.  
 PF 20-SEP-1993; G01974.  
 PR 21-SEP-1992; GB-019936.  
 PA (PITM) PITMAN MOORE INC.  
 PI Francis MJ;  
 DR WPI; 94-118168/14.  
 PT Feline immunodeficiency virus antigenic polypeptide(s) and  
 PT nucleic acid - used to prepare prods. for combating or diagnosis  
 PT of feline immunodeficiency infection  
 PS Disclosure; Figure 2; 68pp; English.  
 CC The sequences given in R51248-R51262 and R58584-85 represent the  
 CC feline immunodeficiency virus (FIV) envelope protein from different  
 CC strains and serotypes. The consensus sequence based on these, is given  
 CC in R51247. The synthetic FIV peptides of the invention were derived  
 CC principally from a combination of the sequences of the FIV UK 8 and  
 CC Petaluma isolates, a composite sequence of which is given in R51246.  
 CC Variations in the sequence may occur between different strains or  
 CC serotypes, isolates of different geographical origin or even between  
 CC different isolates from the same host.  
 SQ Sequence 856 AA;

Query Match 44.4%; Score 36; DB 1; Length 856;  
 Best Local Similarity 66.7%; Pred. NO. 2.2e+02;  
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 6 IKINNIRIV 14  
 :||||:|  
 Db .660 LKINLRV 668

Search completed: February 8, 2000, 04:05:41  
 Job time: 9358 sec

**THIS PAGE BLANK (USPTO)**

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	44	54.3	471	2	F71618	adenylosuccinate 1	
2	43	53.1	265	2	S32735	AAF protein - Esch	
3	41	50.6	364	2	S05434	hypothetical prote	
4	41	50.6	459	2	G71619	PINT domain protei	
5	40	48.4	294	2	H38888	COI inron 15 prot	
6	40	49.4	334	2	A70101	hypothetical prote	
7	40	49.4	579	2	T01160	hypothetical prote	
8	40	49.4	1204	2	S62506	hypothetical prote	
9	39	48.1	275	2	F71714	2-dehydro-3-deoxy	
10	39	48.1	624	2	EA4483	hypothetical prote	
11	38.5	47.5	2875	1	RRVUTW	genome polyprotein	
12	38	46.9	82	2	S07352	hypothetical prote	
13	38	46.9	338	2	F71690	hypothetical prote	
14	38	46.9	373	2	F71636	hypothetical prote	
15	38	46.9	436	2	E70387	conserved hypothet	
16	38	46.9	513	2	A53739	hypothetical vacuo	
17	38	46.9	763	2	S51300	probable membrane	
18	38	46.9	801	2	D70309	ribonucleoside-dip	
19	38	46.9	835	2	F70363	cation transportin	
20	38	46.9	4572	2	S57908	hypothetical 527K	
21	37.5	46.3	354	2	A69877	anion permease hom	
22	37	45.7	197	1	Y2BST	tunicamycin resist	
23	37	45.7	254	2	S23475	rod-core linker po	
24	37	45.7	320	2	T03285	hypothetical prote	
25	37	45.7	325	2	B72475	probable transcrip	
26	37	45.7	543	2	I40545	oligopeptide ABC t	
27	37	45.7	1487	2	S62048	probable membrane	
28	36	44.4	91	2	F70142	glu-tRNA amidotran	
29	36	44.4	178	2	B71942	glucose inhibited	
30	36	44.4	235	1	S0MS	parotid secretory	

3 KLVKINNIRI 13  
Ov

Db 9 KEIKINNIRI 19  
|:|||||

## RESULT 3

S05434

hypothetical protein A3 - Emericella nidulans mitochondrion (SGC3)  
C:Species: mitochondrion Emericella nidulans, Aspergillus nidulans  
C:Date: 01-Dec-1989 #sequence\_revision 01-Dec-1989 #text\_change 09-Sep-1997  
C:Accession: S05434  
R:Brown, T.A.; Constable, A.; Waring, R.B.; Scazzocchio, C.; Davies, R.W.  
Nucleic Acids Res. 17, 5838, 1989  
A:Title: Nucleotide sequence of the only unidentified reading frame in the Aspergillus n

A:Reference number: S05434; MUID:89345173

A:Accession: S05434

A:Molecule type: DNA

A:Residues: 1-364 &lt;BRO&gt;

A:Cross-references: EMBL:X15442; NID:gl2720; PID:gl2721

C:Genetics:

A:Genome: mitochondrion

A:Genetic code: SGC3

C:Keywords: mitochondrion

Query Match

Best Local Similarity 50.6%; Score 41; DB 2; Length 364;

Matches 6; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 3 KLVIKINNIRIVKVF 17

|:|||||

Db 71 KIIRINTVKIITIF 85

## RESULT 4

G71619

P1NT domain protein (proteasomal subunit) PFB0240w - malaria parasite (Plasmodium falciparum)  
C:Species: Plasmodium falciparum

C:Date: 13-Nov-1998 #sequence\_revision 13-Nov-1998 #text\_change 07-May-1999

C:Accession: G71619

R:Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.V.;

; Pertea, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.;

Science 282, 1128-1132, 1998

A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.

A:Reference number: A71600; MUID:99021743

A:Accession: G71619

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-459 &lt;GAR&gt;

A:Cross-references: GB:AE001382; GB:AE001362; NID:g3845130; PID:g3845132; TIGR:PFB0240w

A:Experimental source: clone 3D7

C:Genetics:

A:Gene: PFB0240w

Query Match

Best Local Similarity 50.6%; Score 41; DB 2; Length 459;

Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 SYKLVIKINNIRIVKVF 17

|:|||||

Db 219 SIKLCMKLNMQITTSF 235

## RESULT 5

H38888

COI intron 15 protein - Podospora anserina mitochondrion (SGC3)

C:Species: mitochondrion Podospora anserina

C:Date: 03-Feb-1994 #sequence\_revision 03-Feb-1994 #text\_change 09-Sep-1997

C:Accession: H38888

R:Cummings, D.J.; Michel, F.; McNally, K.L.

Curr. Genet. 16, 381-406, 1989

A:Title: DNA sequence analysis of the 24.5 kilobase pair cytochrome oxidase subunit I m

A:Reference number: A48327; MUID:90124722

A:Accession: H38888

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-294 &lt;CDM&gt;

A:Cross-references: GB:X55026; NID:gl4030; PID:e33710; PID:gl334548; GB:M30937; GB:M6

C:Genetics:

A:Genome: mitochondrion

A:Genetic code: SGC3

C:Keywords: mitochondrion

Query Match

Best Local Similarity 49.4%; Score 40; DB 2; Length 294;

Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 YKLVIKINNIR 12

|:|||||

Db 91 YRLVIKLSNIQ 101

## RESULT 6

A70101

hypothetical protein BB0009 - Lyme disease spirochete

C:Species: Borrelia burgdorferi (Lyme disease spirochete)

C:Date: 13-Feb-1998 #sequence\_revision 13-Feb-1998 #text\_change 05-Jun-1998

C:Accession: A70101

R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Wh

son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vu

; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.

Nature 390, 580-586, 1997

A:Authors: Smith, H.O.; Venter, J.C.

A:Title: Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.

A:Reference number: A70100; MUID:98065943

A:Accession: A70101

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-334 &lt;KLE&gt;

A:Cross-references: GB:AE001115; GB:AE000783; NID:g2687879; PID:g2687892; TIGR:BB0009

A:Experimental source: strain B31

Query Match

Best Local Similarity 49.4%; Score 40; DB 2; Length 334;

Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 SYKLVIKINNIRIV 14

|:|||||

Db 101 SYKLPIKIKLNISI 114

## RESULT 7

T01160

hypothetical protein F7N22.6 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 14-May-1999

C:Accession: T01160

R:Dante, M.

submitted to the EMBL Data Library, April 1998

A:Description: The sequence of A. thaliana F7N22.

A:Reference number: Z14250

A:Accession: T01160

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-579 &lt;DAN&gt;

A:Cross-references: EMBL:AF058825; NID:g3047060; PID:g3047066

A:Experimental source: cultivar Columbia

C:Genetics:

A:Map position: 4

A:Introns: 70/2; 380/1

A&gt;Note: F7N22.6

Query Match

49.4%; Score 40; DB 2; Length 579;



genome polyprotein tomato spotted wilt virus (strain SR-02)  
N; Alternate names: L protein  
C; Contains: RNA-directed RNA polymerase (EC 2.7.7.48)  
C; Species: tomato spotted wilt virus  
C; Date: 30-Sep-1992 #sequence\_revision 30-sep-1992 #text\_change 29-May-1998  
C; Accession: JQ1335  
R; de Haan, P.; Kormelink, R.; de Oliveira Resende, R.; van Poelwijk, F.; Peters, D.;  
T. Gen. Virol. 71: 2207-2216, 1991

A; Reference number: JQ1335  
A; Accession: JQ1335  
A; Molecule type: genomic RNA  
A; Residues: 1-2875 <DEH>  
C; Genetics:  
A; Map position: segment L  
C; Superfamily: bunyavirus RNA-directed RNA polymerase  
C; Keywords: nucleotidyltransferase

hypocotylar protein 82, faba bean chloroplast  
C:Species: chloroplast *Vicia faba* (faba bean)  
C:date: 02-Dec-1993 #sequence\_revision 01-Dec-1995 #text\_change 09-Sep-1997  
C:Accession: S07352  
R:Kuntz, M.; weil, J.H.; Steinmetz, A.  
Nucleic Acids Res. 12, 5037-5047, 1984  
A:Title: Nucleotide sequence of a 2 kbp BamHI I fragment of *Vicia faba* chloroplast DNA  
A:Reference number: S07352; MID:84247356

A;Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-82 <RUN>  
A:Cross-references: EMBL:X00682; NID:g12385; PID:g12386  
A>Note: the authors translated the codon CAA for residue 51 as Glu  
C;Keywords: chloroplast

Query Match 46.9%; Score 38; DB 2; Length 82;

Best Local Similarity 58.3%; Pred. No. 9.1;  
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 YKLVKINNRI 13  
:||||:|:|  
Db 35 YRLVIGNNLEI 46

RESULT 13  
F71690  
hypothetical protein RP340 - Rickettsia prowazekii  
C:Species: Rickettsia prowazekii  
C:Date: 21-Nov-1998 #sequence\_revision 21-Nov-1998 #text\_change 07-May-1999  
C:Accession: F71690  
R:Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sichert-Ponten, T.; Alsmark,  
Nature 396, 133-140, 1998  
A:Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.  
A:Reference number: A71630; MUID:99039499  
A:Accession: F71690  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-338 <AND>  
A:Cross-references: GB:AJ235271; GB:AJ235269; NID:g3860788; PID:e1342644; PID:g3860900  
A:Experimental source: strain Madrid E  
C:Genetics:  
A:Gene: RP340

Query Match 46.9%; Score 38; DB 2; Length 338;  
Best Local Similarity 50.0%; Pred. No. 39;  
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 3 KLVIKINNRIIVK 16  
:|:|:|:|:|  
Db 132 RLGLKFNIRIIVQ 145

RESULT 14  
D71636  
hypothetical protein RP760 - Rickettsia prowazekii  
C:Species: Rickettsia prowazekii  
C:Date: 21-Nov-1998 #sequence\_revision 21-Nov-1998 #text\_change 07-May-1999  
C:Accession: D71636  
R:Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sichert-Ponten, T.; Alsmark,  
Nature 396, 133-140, 1998  
A:Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.  
A:Reference number: A71630; MUID:99039499  
A:Accession: D71636  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-373 <AND>  
A:Cross-references: GB:AJ235273; GB:AJ235269; NID:g3861237; PID:e1343033; PID:g3861289  
A:Experimental source: strain Madrid E  
C:Genetics:  
A:Gene: RP760

Query Match 46.9%; Score 38; DB 2; Length 373;  
Best Local Similarity 50.0%; Pred. No. 43;  
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 5 VIKINNRIIVK 16  
:||||:|:|  
Db 244 IIKNNLNDIIR 255

RESULT 15  
E70387  
conserved hypothetical protein aq\_1013 - Aquifex aeolicus  
C:Species: Aquifex aeolicus  
C:Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 08-May-1998  
C:Accession: E70387  
R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ov

V.  
Nature 392, 353-358, 1998  
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.  
A:Reference number: A70300; MUID:98196666  
A:Accession: E70387  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-436 <AQF>  
A:Cross-references: GB:AE000718; NID:g2983504; PID:g2983516; GB:AE000657  
A:Experimental source: strain VFS  
C:Genetics:  
A:Gene: aq\_1013

Query Match 46.9%; Score 38; DB 2; Length 436;  
Best Local Similarity 46.7%; Pred. No. 50;  
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 2 YKLVKINNRIIVK 16  
:|:|:|:|:|  
Db 125 YKSLIYLTNIEVLK 139

Search completed: February 7, 2000, 18:04:36  
Job time: 22202 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2000, 01:25:56 ; Search time 68.63 Seconds  
(without alignments)  
7.398 Million cell updates/sec

Title: US-08-653-294-22

Perfect score: 81

Sequence: 1 SYKLVIKINNIRIVVKF 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 82229 seqs, 29864866 residues

Total number of hits satisfying chosen parameters: 82229

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : SwissProt\_38.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	43	53.1	265	1 AGGR_ECOLI	P43464 escherichia
2	40	49.4	334	1 Y009_BORBU	O51042 borrelia bu
3	40	49.4	1204	1 YAEF_SCHPO	Q09854 schizosacch
4	39	48.1	275	1 KDSA_RICPR	Q9ze84 rickettsia
5	39	48.1	624	1 YE70_METJA	Q58865 methanococc
6	38.5	47.5	2875	1 RRPL_TSWV1	P28976 tomato spot
7	38	46.9	82	1 YC11_VICFA	P08889 vicia faba
8	38	46.9	763	1 YN51_YEAST	P42843 saccharomyc
9	38	46.9	801	1 RIR1_AQUAE	O65503 aquifex aeo
10	38	46.9	1310	1 VAC3_HELPY	Q48253 helicobacte
11	37	45.7	197	1 TMRB_BACSU	P12921 bacillus su
12	37	45.7	253	1 PYG3_NASLA	P29733 mastigoclad
13	37	45.7	284	1 SMLA_DICDI	P54661 dictyosteli
14	37	45.7	412	1 RNC_CAEEL	O01326 caenorhabdi
15	37	45.7	543	1 APPA_BACSU	P42061 bacillus su
16	37	45.7	546	1 YAO5_SCHPO	Q10084 schizosacch
17	37	45.7	1487	1 MDS3_YEAST	P53094 saccharomyc
18	36	44.4	91	1 GATC_BORBU	O51318 borrelia bu
19	36	44.4	235	1 PSP_MOUSE	P07743 mus musculu
20	36	44.4	240	1 PYRH_HELPY	P56106 helicobacte
21	36	44.4	241	1 FRBY_HAEIN	P47760 haemophilus
22	36	44.4	352	1 DHG_THEAC	P13203 thermoplasm
23	36	44.4	427	1 THAI_BRARE	Q98867 brachydanio
24	36	44.4	515	1 COPD_CAEEL	O09236 caenorhabdi
25	36	44.4	808	1 MCM3_HUMAN	P25205 homo sapien
26	36	44.4	812	1 MCM3_MOUSE	P25206 mus musculu
27	36	44.4	854	1 ENV_FIVSD	P19030 feline immu
28	36	44.4	854	1 ENV_FIVSD	Q05312 feline immu
29	36	44.4	855	1 ENV_FIVU8	Q04995 feline immu
30	36	44.4	856	1 ENV_FIVPE	P16090 feline immu
31	36	44.4	862	1 MUT5_BORBU	O51737 borrelia bu
32	36	44.4	918	1 PEP3_YEAST	P27801 saccharomyc
33	36	44.4	1047	1 MSH3_YEAST	P25336 saccharomyc
34	36	44.4	1287	1 VAC2_HELPY	Q48245 helicobacte

35 36 44.4 1296 1 VAC1\_HELPY  
36 35.5 43.8 814 1 PHSG\_CHLFR  
37 35 43.2 180 1 Y088\_METJA  
38 35 43.2 184 1 Y807\_METJA  
39 35 43.2 195 1 NUC\_BORBU  
40 35 43.2 252 1 PYG4\_ANASP  
41 35 43.2 265 1 CFAD\_ECOLI  
42 35 43.2 301 1 Y664\_METJA  
43 35 43.2 374 1 RT04\_ACACA  
44 35 43.2 380 1 RPOA\_CHLFR  
45 35 43.2 435 1 KICH\_RAT

Q48247 helicobacte  
O84250 chlamydia t  
Q57553 methanococc  
Q58217 methanococc  
Q51372 borrelia bu  
P29989 anabaena sp  
P25393 escherichia  
Q58078 methanococc  
P45755 acanthamoeb  
Q46449 chlamydia t  
Q01134 rattus norv

#### ALIGNMENTS

RESULT 1  
AGGR\_ECOLI  
ID AGGR\_ECOLI STANDARD; PRT: 265 AA.  
AC P43464;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 01-NOV-1995 (Rel. 32, Last annotation update)  
DE TRANSCRIPTIONAL ACTIVATOR AGGR (AAF/I REGULATORY PROTEIN).  
GN AGGR.  
OS Escherichia coli.  
OG Plasmid.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ENTEROAGGREGATIVE 17-2 / SEROTYPE 03:H2;  
RX MEDLINE: 94321342.  
RA NATARO J.P., YIKANG D., YINGKANG D., WALKER K.;  
RT "Aggr, a transcriptional activator of aggregative adherence fimbria I  
expression in enteroaggregative Escherichia coli.";  
RL J. Bacteriol. 176:4691-4699(1994).  
CC -!- FUNCTION: TRANSCRIPTIONAL ACTIVATOR OF AGGREGATIVE ADHERENCE  
FIMBRIA I EXPRESSION IN ENTEROAGGREGATIVE ESCHERICHIA COLI.  
CC -!- SIMILARITY: BELONGS TO THE ARAC/XYLS FAMILY OF TRANSCRIPTIONAL  
REGULATORS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: Z32523; CAA83535.1; .  
DR EMBL: Z18751; CAA79242.1; .  
DR PROSITE: PS00041; HTH\_ARAC\_FAMILY\_1; 1.  
DR PROSITE: PS01124; HTH\_ARAC\_FAMILY\_2; 1.  
DR PFAM: PF00165; HTH\_2; 1.  
KW Transcription regulation; Activator; DNA-binding; Plasmid.  
FT DNA\_BIND 180 199 H-T-H MOTIF (BY SIMILARITY).  
SQ SEQUENCE 265 AA; 30689 MW; 480C8676 CRC32;

Query Match 53.1%; Score 43; DB 1; Length 265;  
Best Local Similarity 81.8%; Pred. No. 2;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 KLVIKINNIRI 13  
Db 9 KEIKINNIRI 19  
| :|||||||

#### RESULT 2

Y009\_BORBU  
ID Y009\_BORBU STANDARD; PRT: 334 AA.  
AC O51042;

```

Query Match      48.1%; Score 39; DB 1; Length 275;
Best Local Similarity 63.8%; Pred. NO. 10;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

3 KLVIKINNRI 13
  | | : | | : |
2 KKVVLNNIKI 12

```

```
RESULT 5
YE70 METJA
ID YE70 METJA STANDARD; PRT; 624 AA.
AC Q58865;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE HYPOTHETICAL PROTEIN MJ1470.
GN MJ1470
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
OC Methanococcus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE; 96337999.
RA BULT C.J., WHITE O., OLSEN G.J., ZHOU L., FLEISCHMANN R.D.,
RA SUTTON G.G., BLAKE J.A., FITZGERALD L.M., CLAYTON R.A., GOCAYNE J.D.,
RA KERLAVAGE A.R., DOUGHERTY B.A., TOMB J.-F., ADAMS M.D., REICH C.I.,
RA OVERBECK R., KIRKNESS E.F., WEINSTOCK K.G., MERRICK J.M., GLODEK A.,
RA SCOTT J.L., GEOGHAGEN N.S.M., WEIDMAN J.F., FUHRMANN J.L., NGUYEN D.,
RA UTTERBACK T.R., KELLEY J.M., PETERSON J.D., SADOW P.W., HANNA M.C.,
RA COTTON M.D., ROBERTS K.M., HURST M.A., RAINE B.P., BORODOVSKIY M.,
RA KLENK H.-P., FRASER C.M., SMITH H.O., WOESE C.R., VENTER J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
jannaschii."
RL Science 273:1058-1073(1996).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U67588; AAB99485.1; -.
KW Hypothetical protein; Transmembrane.
FT TRANSMEM 10 30 POTENTIAL.
SQ SEQUENCE 624 AA; 72857 MW; 4C4411ED CRC32;

Query Match 48.1%; Score 39; DB 1; Length 624;
Best Local Similarity 53.3%; Pred. No. 25;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 2 YKLVKINNIRIVK 16
| | | | |
Db 609 YVLTSPNNVTIVK 623

RESULT 6
RRPL_TSWV1
ID RRPL_TSWV1 STANDARD; PRT; 2875 AA.
AC P28976;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-DEC-1992 (Rel. 24, Last annotation update)
DE RNA-DIRECTED RNA POLYMERASE (EC 2.7.7.46) (L PROTEIN).
GN L.
OS Tomato spotted wilt virus (strain Brazilian Br-01) (TSWV).
OC Viruses; ssRNA negative-strand viruses; Bunyaviridae; Tospovirus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 91374019.
RA DE HAAN P., KORMELINK R., DE OLIVEIRA RESENDE R., VAN POELWIJK F.,
RA PETERS D., GOLDBACH R.;
RT "Tomato spotted wilt virus L RNA encodes a putative RNA polymerase.";
RL J. Gen. Virol. 72:2207-2216(1991).
CC -!- CATALYTIC ACTIVITY: N NUCLEOSIDE TRIPHOSPHATE = N PYROPHOSPHATE
CC + RNA(N).
CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X00682; CAA25290.1; -.
KW PIR; S07352; S07352.
KW Chloroplast; Hypothetical protein.
SQ SEQUENCE 82 AA; 9935 MW; 18728241 CRC32;

Query Match 46.9%; Score 38; DB 1; Length 82;
Best Local Similarity 58.3%; Pred. No. 4.3;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 2 YKLVKINNIRI 13
| | | | |
Db 35 YRLVIGMNNLEI 46

RESULT 8
YNSI_YEAST
ID YNSI_YEAST STANDARD; PRT; 763 AA.
AC P42843;
```

DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE HYPOTHETICAL 88.9 KD PROTEIN IN RFA2-STB1 INTERGENIC REGION.  
 GN YNL311C OR N0376.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Hemiascomycetes; Saccharomycetales;  
 OC Saccharomycetaceae; Saccharomycetes.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-S288C / FY1676;  
 RX MEDLINE; 96076632.  
 RA MAFTAH M., NICAUD J.-M., LEVESQUE H., GAILLARDIN C.;  
 RT "Sequencing analysis of a 24.7 kb fragment of yeast chromosome XIV  
 RT identifies six known genes, a new member of the hexose transporter  
 RT family and ten new open reading frames.";  
 RL Yeast 11:1077-1085(1995).  
 [2]  
 RN SEQUENCE OF 149-763 FROM N.A.  
 RP MAURER C.T.C., URBANUS J.H.M., PLANTA R.J.;  
 RA Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
 RL [1]  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; Z46259; CAA86384.1; -  
 DR EMBL; Z71587; CAA96240.1; -  
 DR PFAM; PF00646; F-box; 1.  
 KW Hypothetical protein.  
 FT DOMAIN 22 28 POLY-GLU.  
 SQ SEQUENCE 763 AA; 88941 MW; F58683F4 CRC32;

Query Match 46.9%; Score 38; DB 1; Length 763;  
 Best Local Similarity 52.6%; Pred. No. 46;  
 Matches 10; Conservative 2; Mismatches 1; Indels 6; Gaps 1;

QY 1 SYKL-----VIKINIRI 13  
 |||| |I:||||  
 Db 177 SYKLDSYDKVKLNIRL 195

RESULT 9  
 RIR1\_AQUAE STANDARD; PRT; 801 AA.  
 AC O66503;  
 DT 15-DEC-1999 (Rel. 39, Created)  
 DT 15-DEC-1999 (Rel. 39, Last sequence update)  
 DT 15-DEC-1999 (Rel. 39, Last annotation update)  
 DE RIBONUCLEOSIDE-DIPHOSPHATE REDUCTASE ALPHA CHAIN (EC 1.17.4.1)  
 DE (RIBONUCLEOTIDE REDUCTASE).  
 GN NRDA OR AQ.094.  
 OS Aquifex aeolicus.  
 OC Bacteria; Aquificales; Aquificaceae; Aquifex.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-VF5;  
 RX MEDLINE; 98196666.  
 RA DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,  
 RA GRAHAM D.E., OVERBEK R., SNEAD M.A., KELLER M., AUWAY M., HUBER R.,  
 RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;  
 RT "The complete genome of the hyperthermophilic bacterium Aquifex  
 RT aeolicus.";  
 RL Nature 392:353-358(1998).  
 CC -!- FUNCTION: CATALYZES THE BIOSYNTHESIS OF DEOXYRIBONUCLEOTIDES FROM  
 CC THE CORRESPONDING RIBONUCLEOTIDES, PRECURSORS THAT ARE NECESSARY  
 CC FOR DNA SYNTHESIS (BY SIMILARITY).  
 CC -!- CATALYTIC ACTIVITY: 2'-DEOXYRIBONUCLEOSIDE DIPHOSPHATE + OXIDIZED

CC THIOREDOXIN + H(2)O - RIBONUCLEOSIDE DIPHOSPHATE + REDUCED  
 CC THIOREDOXIN.  
 CC -!- PATHWAY: FIRST REACTION IN THE DNA REPLICATION PATHWAY.  
 CC -!- SUBUNIT: TETRAMER OF TWO ALPHA AND TWO BETA CHAINS  
 CC (BY SIMILARITY).  
 CC -!- SIMILARITY: BELONGS TO THE RIBONUCLEOSIDE DIPHOSPHATE REDUCTASE  
 CC LARGE CHAIN FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AE000673; AAC08460.1; -  
 DR PROSITE; PS00089; RIBORED\_LARGE; 1.  
 DR PFAM; PF00317; ribonucleo; 3.  
 KW Oxidoreductase; DNA replication.  
 FT ACT\_SITE 235 235 BY SIMILARITY.  
 FT ACT\_SITE 485 485 BY SIMILARITY.  
 FT ACT\_SITE 521 521 BY SIMILARITY.  
 FT SITE 796 796 INTERACTS WITH THIOREDOXIN/THIOREDOXIN  
 FT SITE 799 799 INTERACTS WITH THIOREDOXIN/THIOREDOXIN  
 FT SITE 799 799 INTERACTS WITH THIOREDOXIN/THIOREDOXIN  
 FT SITE 799 799 (BY SIMILARITY).  
 FT SITE 799 799 (BY SIMILARITY).  
 SQ SEQUENCE 801 AA; 92913 MW; 7038CF09 CRC32;

Query Match 46.9%; Score 38; DB 1; Length 801;  
 Best Local Similarity 58.3%; Pred. No. 49;  
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 6 IKINIRIVVKF 17  
 : ||||| :||  
 Db 15 LDINKIRIAIKF 26

RESULT 10  
 VAC3\_HELPHY STANDARD; PRT; 1310 AA.  
 ID VAC3\_HELPHY  
 AC Q48253;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE VACUOLATING CYTOTOXIN PRECURSOR.  
 GN VACA.  
 OS Helicobacter pylori (Campylobacter pylori).  
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
 OC Helicobacter.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-TX30A;  
 RX MEDLINE; 95355366.  
 RA AHERTON J.C., CAO P., PECK R.M. JR., TUMMURU M.K., BLASER M.J.,  
 RA COVER T.L.;  
 RT "Mosaicism in vacuolating cytotoxin alleles of Helicobacter pylori.  
 RT Association of specific vacA types with cytotoxin production and  
 RT peptic ulceration.";  
 RL J. Biol. Chem. 270:17771-17777(1995).  
 CC -!- FUNCTION: INDUCES VACUOLATION OF EUKARYOTIC CELLS. CAUSES  
 CC ULCERATION AND GASTRIC LESIONS.  
 CC -!- SUBCELLULAR LOCATION: SECRETED.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC

```
DR EMBL; U29401; AAA86834.1; -.
KW Cytotoxin; Toxin; Signal.
FT SIGNAL 1 30
FT CHAIN 31 ? VACUOLATING CYTOTOXIN.
FT PROPEP ? 1310 POTENTIAL.
SQ SEQUENCE 1310 AA; 141988 MW; BDE25580 CRC32;

Query Match 46.9%; Score 38; DB 1; Length 1310;
Best Local Similarity 53.3%; Pred. No. 83;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 YKLVKINNIRIVK 16
II :||||| :|
DB 725 YKPLKINNAQLTK 739

RESULT 11
TMRB_BACSU STANDARD; PRT; 197 AA.
AC P12921;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE TUNICAMYCIN RESISTANCE PROTEIN.
GN TMRB.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
RN [1]
RP SEQUENCE FROM N.A.
RA HARADA S., YODA K., MORI M., TANIMOTO A., FURUSATO T., YAMANE K.,
RA TAKATSUKI A., YAMASAKI M., TAMURA G.;
RT "The gene responsible for tunicamycin resistance, tnrB, in Bacillus
RT subtilis.";
RL Agric. Biol. Chem. 52:1785-1789(1988).
RN [2]
RP PARTIAL SEQUENCE OF 1-10, AND CHARACTERIZATION.
RX MEDLINE; 92325013.
RA NODA V., YODA K., TAKATSUKI A., YAMASAKI M.;
RT "TnrB protein, responsible for tunicamycin resistance of Bacillus
RT subtilis, is a novel ATP-binding membrane protein.";
RL J. Bacteriol. 174:4302-4307(1992).
CC -!- FUNCTION: INVOLVED IN THE RESISTANCE TO TUNICAMYCIN. BINDS ATP.
CC -!- SUBCELLULAR LOCATION: MEMBRANE-ASSOCIATED.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D13793; BAA02946.1; -.
DR EMBL; D13793; BAA02944.1; ALT_INIT.
DR EMBL; D13793; BAA02945.1; ALT_INIT.
DR PIR; JT0240; YTEST.
DR SUBTILIS; BG10456; TMRB.
KW Antibiotic resistance; ATP-binding; Membrane.
FT NP_BIND 7 14
FT ATP (POTENTIAL).
SQ SEQUENCE 197 AA; 22730 MW; 264B136D CRC32;

Query Match 45.7%; Score 37; DB 1; Length 197;
Best Local Similarity 54.5%; Pred. No. 16;
Matches 6; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 3 KLVKINNIRI 13
:|:|:|:|:|
DB 186 RLWKLNHIRI 196

EMBL; U29401; AAA86834.1; -.
PYG3_MASLA STANDARD; PRT; 253 AA.
AC P29733;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE PHYCOBILISOME ROD-CORE LINKER POLYPEPTIDE CPCG3 (L-RC 29.6).
GN CPCG3.
OS Mastigocladus laminosus (Fischerella sp.).
OC Bacteria; Cyanobacteria; Stigonematales; Fischerella.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PCG 7603;
RX MEDLINE; 92249337.
RA GLAUSER M., STIREWALT V.L., BRYANT D.A., SIDLER W., ZUBER H.;
RT "Structure of the genes encoding the rod-core linker polypeptides of the
RT Mastigocladus laminosus phycobilisomes and functional aspects of the
RT phycobiliprotein/linker-polypeptide interactions.";
RL Eur. J. Biochem. 205:927-937(1992).
CC -!- FUNCTION: ROD-CORE LINKER PROTEIN REQUIRED FOR ATTACHMENT OF
CC PHYCOCYANIN TO ALLOPHYCOCYANIN IN CORES OF PHYCOBILISOMES.
CC -!- FUNCTION: LINKER POLYPEPTIDES DETERMINE THE STATE OF AGGREGATION
CC AND THE LOCATION OF THE DISC-SHAPED PHYCOBILIPROTEIN UNITS WITHIN
CC THE PHYCOBILISOME AND MODULATE THEIR SPECTROSCOPIC PROPERTIES IN
CC ORDER TO MEDIATE A DIRECTED AND OPTIMAL ENERGY TRANSFER.
CC -!- SUBUNIT: THE PHYCOBILISOME IS A HEMIDISCOIDAL STRUCTURE THAT IS
CC COMPOSED OF TWO DISTINCT SUBSTRUCTURES: A CORE COMPLEX (THAT IS
CC CONTAINS PHYCOBILIPROTEINS) AND A NUMBER OF RODS RADIATING FROM
CC THE CORE.
CC -!- SIMILARITY: TO OTHER PHYCOBILISOME LINKER PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X59763; CAA42435.1; -.
DR PIR; S16060; S16060.
DR PIR; S23475; S23475.
DR PFAM; PF00427; PBS_linker_poly; 1.
KW Phycobilisome; Photosynthesis; Multigene family.
FT INIT_MET 0 BY SIMILARITY.
SQ SEQUENCE 253 AA; 29493 MW; 820A835D CRC32;

Query Match 45.7%; Score 37; DB 1; Length 253;
Best Local Similarity 36.8%; Pred. No. 21;
Matches 7; Conservative 6; Mismatches 2; Indels 4; Gaps 1;

QY 2 YKLVKINNIRIVK 16
I:|:|:|:|:|
DB 94 YRLVSVNNRYRLVDVVLK 112

RESULT 13
SMLA_DICDI STANDARD; PRT; 284 AA.
AC P54661; Q23850;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE SMALL AGGREGATE FORMATION PROTEIN.
GN SMLA.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Dictyostelida; Dictyostelium.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 96379719.
RA BROCK D.A., BUCZYNSKI G., SPANN T.P., WOOD S.A., CARDELLI J.,
```

DR WORMPEP; F26E4.13; CE09697.  
DR PROSITE; PS00517; RIBONUCLEASE\_III; 2.  
DR PFAM; PF00836; Ribonuclease\_3; 2.  
KW Hypothetical protein; Hydrolase; Nuclease; Endonuclease.  
SQ SEQUENCE 412 AA: 47087 MW; B6C1ICD1 CRC32;

Query Match 45.7%; Score 37; DB 1; Length 41  
Best Local Similarity 50.0%; Pred. No. 36;  
Matches 7; Conservative 4; Mismatches 3; Indels

OY 3 KLVKINNIRIVVK 16  
DI :|::|||::|  
DB 168 RGCIQFNRIILAK 181

RESULT 15  
ID APPA\_BAGSU STANDARD; PRT; 543 AA.  
AC P42061.  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DE 01-NOV-1995 (Rel. 32, Last annotation update)  
DE OLIGOPEPTIDE-BINDING PROTEIN APPA PRECURSOR.  
GN APPA.  
OS Bacillus subtilis.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
OC Bacillus/Staphylococcus group; Bacillus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=168;  
RA MEDLINE; 95089678.  
RX KOIDE A.; HOCH J.A.;  
RT "Identification of a second oligopeptide transport system  
RL subtilis and determination of its role in sporulation";  
RL Mol. Microbiol. 13:417-426(1994).  
CC -!- FUNCTION: THIS PROTEIN IS A COMPONENT OF AN OLIGOPEP-  
CC PERMEASE, A BINDING PROTEIN-DEPENDENT TRANSPORT SYSTEM  
CC APP SYSTEM CAN COMPLETELY SUBSTITUTE FOR THE OPP SYS-  
CC SPORULATION AND GENETIC COMPETENCE. APPA CAN BIND AN  
CC TETRA- AND PENTAPEPTIDES BUT NOT TRIPEPTIDES.  
CC -!- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A  
CC (PROBABLE).  
CC -!- SIMILARITY: BELONGS TO THE BACTERIAL EXTRACELLULAR S-  
CC PROTEIN FAMILY 5.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced throug-  
CC between the Swiss Institute of Bioinformatics and the  
CC the European Bioinformatics Institute. There are no re-  
CC use by non-profit institutions as long as its conten-  
CC modified and this statement is not removed. Usage by  
CC entities requires a license agreement (See http://www.is-  
CC or send an email to license@sib-sib.ch).  
CC -----  
DR EMBL; U20909; AAA62358.1; -.  
DR EMBL; Z99110; CAB12995.1; -.  
DR SWISSTLST; BG11087; APPA.  
DR PROSITE; PS00013; PROKARY\_LIPOPROTEIN; 1.  
DR PROSITE; PS01040; SBP\_BACTERIAL\_5; 1.  
DR PFAM; PF00496; SBP\_bac\_5; 1.  
KW Peptide transport; Transport; Membrane; Signal; Sporulat-  
KW Lipoprotein.  
FT SIGNAL 1 23 POTENTIAL.  
FT CHAIN 24 543 OLIGOPEPTIDE-BINDING PROTEIN.  
FT LIPID 24 24 N-ACYL DIGLYCERIDE (PROBABLY  
SQ SEQUENCE 543 AA: 61947 MW; C52C01CE CRC32;

Query Match 45.7%; Score 37; DB 1; Length 54  
Best Local Similarity 53.8%; Pred. No. 48;  
Matches 7; Conservative 2; Mismatches 4; Indels

OY 2 YKLVIKINNIRIV 14



Db 270 YKTAEKFNKIV 282

Search completed: February 8, 2000, 01:25:57  
Job time: 1557 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2000, 19:16:09 ; Search time 176.54 Seconds  
(without alignments)  
6.677 Million cell updates/sec

Title: US-08-653-294-22  
Perfect score: 81  
Sequence: 1 SYKLVIKINNIRIVVKF 17

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 225878 seqs, 69334122 residues

Total number of hits satisfying chosen parameters: 225878

Minimum DB seq length: 0  
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database : SPTRMBL12.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phage.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	47	58.0	475	2 O32315	O32315 bacillus th
2	44	54.3	471	5 O15918	O15918 plasmidium
3	41	50.6	240	2 O85022	O85022 helicobacte
4	41	50.6	240	2 O9X414	O9X414 helicobacte
5	41	50.6	364	8 O33805	O33805 emericella
6	41	50.6	459	5 O96149	O96149 plasmidium
7	40	49.4	235	2 O85023	O85023 helicobacte
8	40	49.4	253	5 O15869	O15869 plasmidium
9	40	49.4	294	8 O02676	O02676 podospora a
10	40	49.4	579	10 O65223	O65223 arabidopsis
11	40	49.4	584	2 O9XAS5	O9XAS5 staphylococ
12	39	48.1	275	2 O92ER4	O92ER4 rickettsia
13	39	48.1	339	2 O924S2	O924S2 thermotoga
14	39	48.1	547	10 O9XIA4	O9XIA4 arabidopsis
15	39	48.1	727	10 O9ZWC2	O9ZWC2 arabidopsis
16	38	46.9	240	2 O85029	O85029 helicobacte
17	38	46.9	240	2 O9X413	O9X413 helicobacte
18	38	46.9	240	2 O9X410	O9X410 helicobacte
19	38	46.9	248	2 O9ZIE7	O9ZIE7 helicobacte
20	38	46.9	338	2 O9ZDI8	O9ZDI8 rickettsia

21	38	46.9	345	12 O9YW40	O9YW40 melanoplus
22	38	46.9	373	2 O9ZC10	O9ZC10 rickettsia
23	38	46.9	403	2 O86251	O86251 helicobacte
24	38	46.9	403	2 O86252	O86252 helicobacte
25	38	46.9	403	2 O86253	O86253 helicobacte
26	38	46.9	436	2 O67128	O67128 aquifex aeo
27	38	46.9	513	2 O48246	O48246 helicobacte
28	38	46.9	576	11 O62970	O62970 rattus norv
29	38	46.9	835	2 O66938	O66938 aquifex aeo
30	38	46.9	1323	2 O87018	O87018 helicobacte
31	38	46.9	1736	10 O23025	O23025 arabidopsis
32	38	46.9	4572	10 O40712	O40712 oryza sativ
33	37.5	46.3	354	2 O34734	O34734 bacillus su
34	37	45.7	130	2 O44641	O44641 buchnera ap
35	37	45.7	132	12 O66232	O66232 cacao swoll
36	37	45.7	197	2 P94385	P94385 bacillus su
37	37	45.7	240	2 O9X415	O9X415 helicobacte
38	37	45.7	240	2 O9X412	O9X412 helicobacte
39	37	45.7	240	2 O9X409	O9X409 helicobacte
40	37	45.7	320	10 O80756	O80756 arabidopsis
41	37	45.7	325	1 O9Y942	O9Y942 aeropyrum p
42	37	45.7	347	6 O97526	O97526 saimiri bol
43	37	45.7	402	6 O9XS98	O9XS98 cebus apell
44	37	45.7	405	6 O97527	O97527 saimiri sci
45	37	45.7	405	6 O97528	O97528 callithrix

ALIGNMENTS

RESULT 1  
O32315  
ID O32315 PRELIMINARY: PRT: 475 AA.  
AC O32315;  
DT 01-JAN-1998 (TREMBlrel. 05, Created)  
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)  
DE HYPOTHETICAL 53.7 KD PROTEIN.  
OS Bacillus thuringiensis.  
OG Plasmid pGI3.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
OC Bacillus/Staphylococcus group; Bacillus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 97405895.  
RA HOFACK L., SEURINCK J., MAHILLON J.;  
RT "Nucleotide sequence and characterization of the cryptic Bacillus  
RT thuringiensis plasmid pGI3 reveal a new family of rolling circle  
RT replicons."  
RL J. Bacteriol. 179:5000-5008(1997).  
DR EMBL; Y11173; CAA72058.1; -  
KW Hypothetical protein; Plasmid.  
SQ SEQUENCE 475 AA; 53733 MW; FCE7AEC8 CRC32;

Query Match 58.0%; Score 47; DB 2; Length 475;  
Best Local Similarity 50.0%; Pred. No. 5.1;  
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 2 YKLVIKINNIRIVVKF 17  
| | | | | | | | | |  
Db 48 YNLIIKVNMMKDVYSF 63

RESULT 2  
O15918  
ID O15918 PRELIMINARY: PRT: 471 AA.  
AC O15918;  
DT 01-JAN-1998 (TREMBlrel. 05, Created)  
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
DT 01-NOV-1998 (TREMBlrel. 12, Last annotation update)  
DE ADENYLOSUCCINATE LYASE (EC 4.3.2.2).  
GN PFB0295W.

OS Plasmodium falciptarum.  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=B8;  
 RX MEDLINE; 97418807.  
 RA MARSHALL V.M., COPPEL R.L.;  
 RT "Characterisation of the gene encoding adenylosuccinate lyase of  
 Plasmodium falciptarum";  
 RL Mol. Biochem. Parasitol. 88:237-241(1997).  
 [2]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=B8;  
 RX MEDLINE; 98384012.  
 RA MARSHALL V.M., TIEQUAO W., COPPEL R.L.;  
 RT "Close linkage of three merozoite surface protein genes on chromosome  
 2 of Plasmodium falciptarum";  
 RL Mol. Biochem. Parasitol. 94:13-25(1998).  
 [3]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE; 99021743.  
 RA GARDNER M.J., TETTELIN H., CARUCCI D.J., CUMMINGS L.M., ARAVIND L.,  
 KOONIN E.V., SHALLOM S., MASON T., YU K., FUJII C., PEDERSON J.,  
 RA SHEN K., JING J., ASTON C., LAI Z., SCHWARTZ D.C., PERTEA M.,  
 RA SALZBERG S., ZHOU L., SUTTON G.G., CLAYTON R., WHITE O., SMITH H.O.,  
 RA FRASER C.M., ADAMS M.D., VENTER J.C., HOFFMAN S.L.;  
 RT "Chromosome 2 sequence of the human malaria parasite Plasmodium  
 falciptarum";  
 RL Science 282:1126-1132(1998).  
 DR EMBL; AF033037; AAC32788.1; -.  
 DR EMBL; AE001385; AAC71848.1; -.  
 DR PROSITE; PS00163; FUMARATE\_LYASES; 1.  
 DR PFAM; PF00206; lyase\_1; 2.  
 KW Lyase.  
 SQ SEQUENCE 471 AA; 55176 MW; A7EC2F0B CRC32;

Query Match 54.3%; Score 44; DB 5; Length 471;  
 Best Local Similarity 56.2%; Pred. No. 16;  
 Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 YKLVKINNIRIVVKF 17  
 ||:|:|:|:|:|  
 Db 194 YKHINKLNIRIVVKF 209

RESULT 3  
 ID O85022 PRELIMINARY; PRT; 240 AA.  
 AC O85022;  
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
 DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).  
 GN VACA.  
 OS Helicobacter pylori (Campylobacter pylori).  
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
 OC Helicobacter.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=R10A;  
 RX MEDLINE; 98314552.  
 RA PAN Z.J., BERG D.E., VAN DER HULST R.W., SU W.W., RAUDONIKIENE A.,  
 RA XIAO S.D., BARKERT J., TYTGAT G.N., VAN DER ENDE A.;  
 RT "Prevalence of vacuolating cytotoxin production and distribution of  
 distinct vacA alleles in Helicobacter pylori from China.";  
 RL J. Infect. Dis. 178:220-226(1998).  
 DR EMBL; AF035609; AAC31125.1; -.  
 FT NON\_TER 1  
 FT NON\_TER 240  
 SQ SEQUENCE 240 AA; 26722 MW; 6A5AD7A4 CRC32;

Query Match 50.6%; Score 41; DB 2; Length 240;  
 Best Local Similarity 53.3%; Pred. No. 26;  
 Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 YKLVKINNIRIVVK 16  
 ||:|:|:|:|:|  
 Db 176 YKPLIKINNAQDLIK 190

RESULT 4  
 Q9X414  
 ID Q9X414 PRELIMINARY; PRT; 240 AA.  
 AC Q9X414;  
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)  
 DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).  
 GN VACA.  
 OS Helicobacter pylori (Campylobacter pylori).  
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
 OC Helicobacter.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=T-25;  
 RA LIN C.W.;  
 RT "Helicobacter pylori Taiwanese isolates";  
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF091822; AAD23359.1; -.  
 FT NON\_TER 1  
 FT NON\_TER 240  
 SQ SEQUENCE 240 AA; 26608 MW; OD37C8B0 CRC32;

Query Match 50.6%; Score 41; DB 2; Length 240;  
 Best Local Similarity 53.3%; Pred. No. 26;  
 Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 YKLVKINNIRIVVK 16  
 ||:|:|:|:|:|  
 Db 176 YKPLIKINNAQDLIK 190

RESULT 5  
 ID Q33805 PRELIMINARY; PRT; 364 AA.  
 AC Q33805;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
 DE ASPERGILLUS NIDULANS MITOCHONDRIAL DNA FOR URF A3.  
 OS Emericella nidulans (Aspergillus nidulans).  
 OG Mitochondrion.  
 OC Eukaryota; Fungi; Ascomycota; Euascomycetes; Plectomycetes;  
 OC Eurotiiales; Trichocomaceae; Emericella.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 89345173.  
 RA BROWN T.A., CONSTABLE A., WARING R.B., SCAZZOCCIO C., DAVIES R.W.;  
 RT "Nucleotide sequence of the only unidentified reading frame in the  
 Aspergillus nidulans mitochondrial genome.";  
 RL Nucleic Acids Res. 17:5838-5838(1989).  
 DR EMBL; X15442; CAA33482.1; -.  
 KW Mitochondrion.  
 SQ SEQUENCE 364 AA; 42482 MW; 89011CF2 CRC32;

Query Match 50.6%; Score 41; DB 8; Length 364;  
 Best Local Similarity 40.0%; Pred. No. 38;  
 Matches 6; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 3 KLVKINNIRIVVKF 17  
 ||:|:|:|:|:|  
 Db 71 KIIIRINTVKIITIF 85

```

RESULT 6
O96149
ID O96149 PRELIMINARY; PRT: 459 AA.
AC O96149;
DT 01-MAY-1999 (TRENBLrel. 10, Created)
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
DT 01-MAY-1999 (TRENBLrel. 10, Last annotation update)
DE PINT DOMAIN PROTEIN (PROTEASOMAL SUBUNIT).
GN PFB0240W.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 99021743.
RA GARDNER M.J., TETTELIN H., CARUCCI D.J., CUMMINGS L.M., ARAVIND L.,
RA KOONIN E.V., SHALLOM S., MASON T., YU K., FUJII C., PEDERSON J.,
RA SHEN K., JING J., ASTON C., LAI Z., SCHWARTZ D.C., PERTEA M.,
RA SALZBERG S., ZHOU L., SUTTON G.G., CLAYTON R., WHITE O., SMITH H.O.,
RA FRASER C.M., ADAMS M.D., VENTER J.C., HOFFMAN S.L.;
RT "Chromosome 2 sequence of the human malaria parasite Plasmodium
RT falciparum.";
RL Science 282:1126-1132(1998).
DR EMBL: AE001382; AAC71837.1; -.
SQ SEQUENCE 459 AA; 54610 MW; E86BDE0D CRC32;

Query Match 50.6%; Score 41; DB 5; Length 459;
Best Local Similarity 47.1%; Pred. No. 47;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 SYKLVKIKNNIRIVVKF 17
Db 219 SIKLCMKLNMMQITTSF 235

RESULT 7
O85023 PRELIMINARY; PRT: 235 AA.
AC O85023;
DT 01-NOV-1998 (TRENBLrel. 08, Created)
DT 01-NOV-1998 (TRENBLrel. 08, Last sequence update)
DT 01-NOV-1998 (TRENBLrel. 08, Last annotation update)
DE VACUOLATING CYTOTOXIN PRECURSOR (FRAGMENT).
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-R13A;
RX MEDLINE: 98314552.
RA PAN Z.J., BERG D.E., VAN DER HULST R.W., SU W.W., RAUDONIKIENE A.,
RA XIAO S.D., DANKERT J., TYTGAT G.N., VAN DER ENDE A.;
RT "Prevalence of vacuolating cytotoxin production and distribution of
RT distinct vacA alleles in Helicobacter pylori from China.";
RL J. Infect. Dis. 178:220-226(1998).
DR EMBL: AF035610; AAC31126.1; -.
FT NON_TER 1
FT NON_TER 235
SQ SEQUENCE 235 AA; 25955 MW; B2C778AC CRC32;

Query Match 49.4%; Score 40; DB 2; Length 235;
Best Local Similarity 53.3%; Pred. No. 37;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 YKLVKIKNNIRIVVK 16
Db 176 YKPLIKINSARDLIK 190

RESULT 8
O15869 PRELIMINARY; PRT: 253 AA.
AC O15869;
DT 01-JAN-1998 (TRENBLrel. 05, Created)
DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)
DT 01-NOV-1998 (TRENBLrel. 08, Last annotation update)
DE ORNITHINE DECARBOXYLASE (EC 4.1.1.17) (FRAGMENT).
GN ODC.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
RN [1]
RP SEQUENCE FROM N.A.
RA MADHUBALA R.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: L-ORNITHINE - PUTRESCINE + CO(2).
CC -1- COFACTOR: PYRIDOXAL-PHOSPHATE.
DR EMBL: Y14231; CAA74609.1; -.
DR PFAM: PF00278; Orn_DAP_Arg_dec; 1.
KW Lyase.
FT NON_TER 1
FT NON_TER 253
SQ SEQUENCE 253 AA; 29489 MW; 4218E85B CRC32;

Query Match 49.4%; Score 40; DB 5; Length 253;
Best Local Similarity 50.0%; Pred. No. 40;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 4 LVKIKNNIRIVVKF 17
Db 4 LCVKCNDEVIK 17

RESULT 9
Q02676 PRELIMINARY; PRT: 294 AA.
AC Q02676;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TRENBLrel. 08, Last annotation update)
DE HYPOTHETICAL 34.0 KD PROTEIN IN COI INTRON 15 REGION.
OS Podospora anserina.
OG Mitochondrion.
OC Eukaryota; Fungi; Ascomycota; Euascomycetes; Pyrenomycetes;
OC Sordariales; Sordariaceae; Podospora.
RN [1]
RP COMPLETE MITOCHONDRIAL GENOME.
RX STRAIN-S;
RX MEDLINE: 90291512.
RA CUMMINGS D.J., MCNALLY K.L., DOMENICO J.M., MATSUURA E.T.;
RT "The complete DNA sequence of the mitochondrial genome of Podospora
RT anserina.";
RL Curr. Genet. 17:375-402(1990).
DR EMBL: X55026; CAA38793.1; -.
DR PFAM: PF00961; Intron_maturase; 2.
KW Hypothetical protein; Mitochondrion.
SQ SEQUENCE 294 AA; 33970 MW; 387D9E8E CRC32;

Query Match 49.4%; Score 40; DB 8; Length 294;
Best Local Similarity 63.6%; Pred. No. 46;
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 YKLVKIKNNIR 12
Db 91 YRLVIKLSNIQ 101

RESULT 10
O65223 PRELIMINARY; PRT: 579 AA.
AC O65223;

```

DT 01-AUG-1998 (TREMELrel. 07, Created)  
 DT 01-AUG-1998 (TREMELrel. 07, Last sequence update)  
 DT 01-NOV-1999 (TREMELrel. 12, Last annotation update)  
 DE F7N22.6 PROTEIN.  
 GN F7N22.6.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;  
 OC core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;  
 OC Arabidopsis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. COLUMBIA;  
 RA WASHU;  
 RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. COLUMBIA;  
 RA DANTE M.;  
 RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. COLUMBIA;  
 RA WATERSTON R.;  
 RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF058825; AAC13580.1; -;  
 DR PRINTS: PR00939; C3HCZNFINGER.  
 SQ SEQUENCE 579 AA; 66041 MW; 1E97C420 CRC32;

Query Match 49.4%; Score 40; DB 10; Length 579;

Best Local Similarity 35.3%; Pred. No. 85;

Matches 6; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

QY 1 SYKLVIKINIRIVVKF 17

:|::|:|:|:|

Db 426 NYKIMKQNNLIPKF 442

RESULT 11

Q9XAS5 PRELIMINARY; PRT; 584 AA.

AC Q9XAS5;  
 DT 01-NOV-1999 (TREMELrel. 12, Created)  
 DT 01-NOV-1999 (TREMELrel. 12, Last sequence update)  
 DT 01-NOV-1999 (TREMELrel. 12, Last annotation update)  
 DE MAP-7. PROTEIN PRECURSOR.  
 GN MAP-7.  
 OS Staphylococcus aureus.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
 OC Bacillus/Staphylococcus group; Staphylococcus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CLINICAL ISOLATE 7;  
 RA HUSSAIN SHAIKH M., HEILMANN C., PETERS G., HERRMAN M.;  
 RT "Cloning, sequencing and expression of map-7 from Staphylococcus  
 aureus clinical isolate 7."  
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ243790; CAB50820.1; -;  
 KW Signal.  
 FT SIGNAL 1 30 POTENTIAL.  
 FT CHAIN 31 584 POTENTIAL.  
 SQ SEQUENCE 584 AA; 65173 MW; 49AE4DA5 CRC32;

Query Match 49.4%; Score 40; DB 2; Length 584;

Best Local Similarity 62.5%; Pred. No. 86;

Matches 10; Conservative 3; Mismatches 1; Indels 2; Gaps 1;

QY 3 KLVIKI--NNIRIVVK 16

||||| |::|:|

Db 185 KLVIKILENNKSVLK 200

RESULT 12

Q9ZE84 PRELIMINARY; PRT; 275 AA.

AC Q9ZE84;  
 DT 01-MAY-1999 (TREMELrel. 10, Created)  
 DT 01-MAY-1999 (TREMELrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TREMELrel. 10, Last annotation update)  
 DE 2-DEHYDRO-3-DEOXYPHOSPHOCTONATE ALDOLASE (KDSA).  
 GN RP062.  
 OS Rickettsia prowazekii.  
 OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;  
 OC Rickettsiaceae; Rickettsiae; Rickettsia.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MADRID E;  
 RX MEDLINE; 99039499.  
 RA ANDERSSON S.G.E., ZOMORODIPOUR A., ANDERSSON J.O.,  
 RA SICHERTZ-PONTEN T., ALSMARK U.C.M., PODOWSKI R.M., NAEGLUND A.K.,  
 RA ERIKSSON A.S., WINKLER H.H., KURLAND C.G.;  
 RT "The genome sequence of Rickettsia prowazekii and the origin of  
 RT mitochondria".  
 RL Nature 396:133-140(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MADRID E;  
 RA ANDERSSON S.G.E.;  
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ235270; CAA14533.1; -;  
 SQ SEQUENCE 275 AA; 30245 MW; 3E960290 CRC32;

Query Match 48.1%; Score 39; DB 2; Length 275;

Best Local Similarity 63.6%; Pred. No. 62;

Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 KLVIKINNIRI 13

|:|:|:|:|

Db 2 KVVVKNLNKI 12

RESULT 13

Q9Z4S2 PRELIMINARY; PRT; 339 AA.

AC Q9Z4S2;  
 DT 01-MAY-1999 (TREMELrel. 10, Created)  
 DT 01-MAY-1999 (TREMELrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TREMELrel. 10, Last annotation update)  
 DE N-ACETYL-GAMMA-GLUTAMYL-PHOSPHATE REDUCTASE (EC 1.2.1.38).  
 GN ARGC.  
 OS Thermotoga neapolitana.  
 OC Bacteria; Thermotogales; Thermotoga.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NS-E;  
 RA MARC F., WEIGEL P., SANTROT M., SAKANYAN V.;  
 RT "Comparative enzymatic analysis of bi- and monofunctional ornithine  
 RT acetyltransferases from thermophilic microorganisms."  
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NS-E;  
 RA DIMOVA D., WEIGEL P., TAKAHASHI M., MARC F., SAKANYAN V.;  
 RT "Gene regulation at high temperatures : characterization of the  
 RT autoreglatable arginine repressor of the hyperthermophile Thermotoga  
 neapolitana".  
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ009897; CAB38109.1; -;  
 KW Oxidoreductase.  
 SQ SEQUENCE 339 AA; 37911 MW; 697E3D80 CRC32;

Query Match 48.1%; Score 39; DB 2; Length 339;

Best Local Similarity 35.78; Pred. No. 76;  
Matches 5; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 SYKLVIKINNIRIV 14  
| | | : : : | : : |  
Db 79 SYRLAMELQNVKII 92

QY 1 SYKLVIKINNIRIVVKF 17  
| : | : : | | : | | |  
Db 253 SFVLCLRSNNLRGVIF 269

Search completed: February 8, 2000, 19:16:12  
Job time: 21499 sec

RESULT 14  
Q9XIA4 PRELIMINARY; PRT; 547 AA.  
AC Q9XIA4  
DT 01-NOV-1999 (TReMBLrel. 12, Created)  
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)  
DE 01-NOV-1999 (TReMBLrel. 12, Last annotation update)  
GN F13F21.19  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC eudicotyledons; Spermatophyta; Magnoliophyta; eudicotyledons;  
OC core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;  
OC Arabidopsis.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA FEDERSPIEL N.A., PALM C.J., CONWAY A.B., CONN L., HANSEN N.F.,  
RA ALTAFI H., ARAUJO R., HUIZAR L., ROWLEY D., BUEHLER E., DUNN P.,  
RA GONZALEZ A., KREMENETSKAIA I., KIM C., LENZ C., LI J., LIU S.,  
RA LUROS S., SCHWARTZ J., SHINN P., TORIUMI M., VYSOTSKAIA V.S.,  
RA WALKER M., YU G., ECKER J., THEOLOGIS A., DAVIS R.W.,  
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AC007504; AAD43162.1; -  
SQ SEQUENCE 547 AA; 59392 MW; AEC76952 CRC32;

Query Match 48.18; Score 39; DB 10; Length 547;

Best Local Similarity 50.08; Pred. No. 1.2e+02;  
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 SYKLVIKINNIRIVVK 16  
| | | : : : | : : |  
Db 444 SSKLPFIEINGMKVIE 459

RESULT 15  
Q9ZWC2 PRELIMINARY; PRT; 727 AA.  
AC Q9ZWC2  
DT 01-MAY-1999 (TReMBLrel. 10, Created)  
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)  
DE 01-MAY-1999 (TReMBLrel. 10, Last annotation update)  
GN F21M11.4  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC eudicotyledons; Spermatophyta; Magnoliophyta; eudicotyledons;  
OC core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;  
OC Arabidopsis.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. COLUMBIA;  
RA FEDERSPIEL N.A., PALM C.J., CONWAY A.B., CONN L., HANSEN N.F.,  
RA ALTAFI H., ARAUJO R., HUIZAR L., ROWLEY D., BUEHLER E., DUNN P.,  
RA GONZALEZ A., KREMENETSKAIA I., KIM C., LENZ C., LI J., LIU S.,  
RA LUROS S., SCHWARTZ J., SHINN P., TORIUMI M., VYSOTSKAIA V.S.,  
RA WALKER M., YU G., ECKER J., THEOLOGIS A., DAVIS R.W.,  
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AC003027; AAD10669.1; -  
SQ SEQUENCE 727 AA; 81342 MW; 5316180B CRC32;

Query Match 48.18; Score 39; DB 10; Length 727;

Best Local Similarity 47.18; Pred. No. 1.5e+02;  
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

**THIS PAGE BLANK (USPTO)**



OM of: US-08-653-294-22 to: GenEmbl.\* out\_format : pfs  
Date: Feb 8, 2000 10:21 PM  
About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 Compugen Ltd.

## Command line parameters:

-MODEL=framet.p2n.model -DEV=xlp  
-O=Cgpl\_1/USPTO\_spool/US08653294/runat\_04022000\_160701\_15779/app\_query.fasta.2  
-DB=GenEmbl -OFMT=fastap -SUFFIX=rge -GAPOP=12.000 -GAPEXT=4.000  
-MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -QGAPOP=4.500  
-QGAPEXT=0.050 -XGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000  
-XGAPEXT=7.000 -START=1 -MATRIX=blosom62 -TRANS=human40.cdi  
-DELEX=7.000 -DOALIGN=200 -THR SCORE=pct -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pfs -NORM=ext -MINLEN=0 -MAXLEN=100000 -USER=US08653294  
-NCPD=6 -ICPU=3 -NO\_XLPXY -WAIT -THREADS=1

## Search information block:

Query: US-08-653-294-22

Query length: 17

Database: GenEmbl.\*

Database sequences: 821193

Database length: -1518192014

Search time (sec): 10176.920000

## score\_list:

Sequence	Strd	Orig	ZScore	EScore	Len	Documentation
gb_pl1:AP000399	-	50.00	102.05	1.2e+03	154180	Oryza sativa genomic
gb_hg3:AC009186	+	49.00	107.53	575.94	48600	Oryza sativa genomic
gb_hg4:AC010873	-	49.00	98.26	1.9e+03	172740	Oryza sativa genomic
gb_hg3:AC011386	-	49.00	96.58	2.3e+03	217175	Oryza sativa genomic
gb_in2:AF021084	-	48.00	138.56	10.76	465	AF021084 Dianous nitidulus cyt
gb_in1:MSQNCAR	-	48.00	112.94	287.55	15455	L04272 Anopheles quadrimacul
gb_hg4:AC012417	-	48.00	103.53	962.00	56025	Oryza sativa genomic
em_in:AC004375	-	48.00	100.28	1.5e+03	87325	Oryza sativa genomic
gb_in2:AC007765	-	48.00	95.70	2.6e+03	163403	Oryza sativa genomic
gb_pl2:CNS019CV	-	47.00	131.82	25.55	780	AL111479 Botrytis cinerea strai
gb_ba1:BTG3REP	+	47.00	112.23	315.09	11365	Y11173 B.thuringiensis plasmi
gb_hg4:AC010688	+	47.00	97.88	2.0e+03	80958	AC010688 Drosophila melanoga
gb_pl1:AB006598	-	47.00	97.56	2.1e+03	84510	AB006598 Arabidopsis thaliana
gb_hg4:AC008308	-	47.00	93.30	3.6e+03	151373	Oryza sativa genomic
gb_pl2:HSAC02064	-	47.00	93.07	3.7e+03	156214	AC002064 Human BAC clone RGO
gb_hg5:AC013460	-	47.00	92.74	3.8e+03	163346	Oryza sativa genomic
gb_hg7:AC009555	-	47.00	92.68	3.9e+03	164713	Oryza sativa genomic
gb_hg4:AC010859	-	47.00	90.45	5.1e+03	223594	AC010859 Homo sapiens clone
gb_hg2:AC006889	-	47.00	89.15	6.1e+03	267118	AC006889 Caenorhabditis eleg
gb_v1:PHV1VGB	-	46.00	118.84	134.96	3069	Z68147 Phocid herpesvirus type
gb_pl2:ATF1111	+	46.00	93.14	3.6e+03	103150	AL079347 Arabidopsis thalian
gb_pl1:ATT22BA	-	46.00	92.77	3.8e+03	108598	AL049876 Arabidopsis thalian
gb_hg2:AC008311	-	46.00	92.67	3.9e+03	109826	AC008311 Drosophila melanoga
gb_hg2:AC007194_3	-	46.00	92.67	3.9e+03	110000	Continuation (4 of 5) of AC0
gb_hg4:AC007488	-	46.00	90.58	5.1e+03	146443	AC007488 Homo sapiens 3q27 B
gb_hg1:AP000457	-	46.00	90.35	5.2e+03	151210	AP000457 Homo sapiens chromo
gb_hg4:AC011177	-	46.00	89.55	5.8e+03	168660	AC011177 Homo sapiens clone
gb_hg1:AP000811	-	46.00	88.97	6.2e+03	182508	AP000811 Homo sapiens clone
gb_pr4:AC007364	-	46.00	88.27	6.8e+03	200908	AC007364 Homo sapiens clone
gb_hg1:PFMALP2	+	46.00	87.09	7.9e+03	235956	AL035475 Plasmodium falcipar
gb_sts:G53007	+	45.00	134.79	17.45	231	G53007 SHGC-81794 Human Homo sa
gb_sts:G53006	+	45.00	134.48	18.15	241	G53006 SHGC-81793 Human Homo sa
gb_ov:GGRNACP49	+	45.00	120.82	104.66	1561	X84806 G.gallus mRNA for CP49.
gb_pr4:AF137131	+	45.00	108.04	539.34	8970	AF137131 Homo sapiens downstre
gb_ba2:AE001116	-	45.00	107.27	595.14	9363	AE001116 Borrelia burgdorferi
gb_in2:MSQTCG	-	45.00	104.11	893.27	15363	L20934 Anopheles gambiae comp
gb_v1:MHVGENE1	-	45.00	101.55	1.2e+03	21798	M55148 Murine coronavirus ope
gb_ba2:AF036485	-	45.00	96.61	2.3e+03	42810	AF036485 Plasmodium falcipar
gb_in1:CELB0563	+	45.00	96.45	2.4e+03	43801	U28740 Caenorhabditis elegans
gb_in2:AF160864	+	45.00	95.88	2.6e+03	47296	AF160864 Tetrahymena pyriform
gb_pl2:ATAC007584	+	45.00	91.63	4.4e+03	84592	AC007584 Arabidopsis thaliana
gb_pl2:ATAC006918	+	45.00	90.10	5.4e+03	104292	AC006918 Arabidopsis thalian

gb\_hg2:AC008209 - 45.00 88.84 6.3e+03 123872 ! AC008209 Drosophila melan  
gb\_hg5:AC010836 + 45.00 86.81 8.2e+03 163667 ! AC010836 Homo sapiens clo  
gb\_hg5:AC012594 + 45.00 86.75 8.3e+03 164997 ! AC012594 Homo sapiens clo

seq\_name: gb\_pl1:AP000399

## seq\_documentation\_block:

LOCUS AP000399 154180 bp DNA PLN 03-DEC-1999  
DEFINITION Oryza sativa genomic DNA, chromosome 6, clone:P0535G04.  
ACCESSION AP000399  
VERSION AP000399.1 GI:5803242  
KEYWORDS  
SOURCE  
ORGANISM

Oryza sativa (cultivar:Nipponbare) DNA, clone:P0535G04.

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales;  
Poaceae; Oryza.

1 (bases 1 to 154180)

Sasaki,T., Matsumoto,T. and Yamamoto,K.

Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, PAC  
clone:P0535G04

Published Only in DataBase (1999) In press

2 (bases 1 to 154180)

Sasaki,T., Matsumoto,T. and Yamamoto,K.

Direct Submission

JOURNAL

AUTHORS

TITLE

JOURNAL

COMMENT

The orientation of the sequence is from T7 to SP6 of the PAC clone.  
Genes were predicted from the integrated results of the  
following:GENSCAN1.0, BLASTX2.0, as well as  
SplicedPredictor (October1998 version). The genomic sequence was  
searched against the non-redundant database NRP (PIR,SWISSPROT,  
GENPEPT, PDB) from MAFF DNABank and the cDNA sequence database at  
RGP. Protein similarities of the coding regions were searched  
against NRP with BLASTP2.0. ESTs represent the identified cDNA  
sequences using BLASTN1.4 with the corresponding DBJ accession no.  
and RFP clone ID.

## FEATURES

Location/Qualifiers  
1..154180  
/organism="Oryza sativa"  
/cultivar="Nipponbare"  
/db\_xref="taxon:4530"  
/chromosome="6"  
/clone="P0535G04"  
complement(10850..10856)  
/note="ESTs AU078033(R0160),D23790(R0160) correspond to a  
region of the predicted gene; hypothetical protein"  
/codon\_start=1  
/protein\_id="BAA83553.1"  
/db\_xref="GI:5803243"  
/translation="MLIEVGIMLLPFTLLVAPVRRMVLVQELQRLMVVACDRRR  
GGRPTEGVLSRLDRDSATIV"  
complement(join(13526..14940,15018..15165))  
/note="similar to hexose carrier protein HEA6 sRCCHCP\_1  
(Q07423)"  
/codon\_start=1  
/protein\_id="BAA83554.1"  
/db\_xref="GI:5803244"  
/translation="MAGVVGAVGESQERRGGAGTGRVTAFLVLSVCVTAGGGVIFGY  
DIGTAGVSSNEFLRFFFEVHRMEGDVRSYCFDSQLLTAFTSSLYVAGLLTT  
FAASRVAGRGRRPMLLGAAGVAGSVDITMVLIGRLVGLGGLGFAQAVP  
LYISEMAPRMRGAFNGFVGVGALAANVINGTEKIRGGWGVSLAAAPVAG  
LLTIGALFPLETNSLIQQQKVERCDVQLLKIRGADDDVADDELDTVAANSAPVAG  
GGGLMLLORRVPOLAMAVMIPFOQVGTGINAIFYAPVLLRTIGMGESASLSAV  
VTGVVGAVLITMFEADRGRTFLFAGGQMLASQVLGGINAALGDDGGVSRW  
AAALILIAVAGGWSGVLGVPSEVFPLEVRSAQSQTVAISFTVFVFAQAF  
LAMLCRRAGVPIEFAWLAAMTAVFYLLLPETKGVPIEEVAGVWRGHWFSRVVGG  
GEEDERNNGKL"  
complement(join(16352..16617,17437..17488,18513..19025))

CDS

/note="hypothetical protein"  
 /codon\_start=1  
 /protein\_id="BAA83555.1"  
 /db\_xref="GI:58032245"  
 /translation="MSIHRAHFAASPIRGSLASAAVSPSAAADALRLSLDAGAGA  
 ADAHPHPLHSKILFFRRGRPLARSYDPPPPAAAAAPPPPPAWRLAWLPAPV  
 DVPDAFVLGAEHGEKEAAAYWADVSESDGEGAGDSAFVDLRLKAVDMDRDK  
 DAMGLAIAGHPGESLEEARVRETWEETDAQWHSREDVKKALTFAYEKAORTNALKV  
 NOICKGVKRRQISADLKIESEPPAFMFPVPGPYAIAHLLISSWAFEGAKAPSSFSNL  
 "  
 complement(join(21523..21536,22175..22245,22419..22771))  
 /notes="ESTs AU077984(C62969),D41648(S4289),  
 AU063584(C62969) correspond to a region of the predicted  
 gene: similar to putative Ser/Thr protein kinase  
 (AC003105)"  
 /codon\_start=1  
 /protein\_id="BAA83556.1"  
 /db\_xref="GI:5803246"  
 /translation="MNIQRKPGDMNCKSQHNLFRSDYCORCHTQPDQLPLDGYVP  
 GGLTSLDIRPGWCYCNCGVNFASRSCFKCGAIVKDLDPAGOGGVANGDFARALDS  
 SAVRQKAGDWICTRPGCNVHNFASRIEYECNAPREAGNVK"  
 complement(join(29617..29738,29858..29936,30110..30171,  
 30324..30396))  
 /notes="similar to OsenOD93a gene for early nodulin  
 (AB018375)"  
 /codon\_start=1  
 /protein\_id="BAA83557.1"  
 /db\_xref="GI:5803247"  
 /translation="MKNKATDSQLIEQKFAKQSHAEATAGLAKAAVAVTCAIP  
 TFAVSRRSKWANLGNPVAKTVIITTAAGMAYFIAGERKKVLSLRKHSLEAKARAK  
 SEYTLTKSN"  
 join(30628..30744,32792..32887)  
 /notes="hypothetical protein"  
 /codon\_start=1  
 /protein\_id="BAA83558.1"  
 /db\_xref="GI:5803248"  
 /translation="MAATKPRAPGARTRFRITPAEAQTSSGAGCNKSSPIVI  
 WEPKGFISDVADGCLLCHNGHRGR"  
 complement(join(34652..34873,34991..35052,35204..35270))  
 /notes="EST C97982(C0324) corresponds to a region of the  
 predicted gene: similar to OsenOD93a gene for early  
 nodulin (AB018375)"  
 /codon\_start=1  
 /protein\_id="BAA83559.1"  
 /db\_xref="GI:5803249"  
 /translation="MATTVTRAHLDOALAKRCSREANLAGVKAATVATIASAVPTL  
 ASVRVPWAKANINPTGQALIICTAAGMAYFVAADKKILSLARRHSFENAPHLKNTS  
 FQGTGRPHPAFFRP"  
 complement(join(38870..39091,39193..39254,39364..39427))  
 /notes="EST AU077941(C12908) corresponds to a region of the  
 predicted gene: similar to OsenOD93a gene for early  
 nodulin (AB018375)"  
 /codon\_start=1  
 /protein\_id="BAA83560.1"  
 /db\_xref="GI:5803250"  
 /translation="MSTVTRAYLDOKLALAKRCSREATLAGAKAAVATVASAVPTLA  
 SVRMLPWAKANINPTGQALIICTAAGMAYFVAADKKILSLARRHSFENAPHLKNTSF  
 QGTGRPHPAFFRP"  
 complement(join(43261..43431,43494..43544,43798..44050,  
 44305..44333))  
 /notes="hypothetical protein"  
 /codon\_start=1  
 /protein\_id="BAA83561.1"  
 /db\_xref="GI:5803251"  
 /translation="MNTNSYSTRNGGGGGDEAENGSGGGGWAAGAAATAVGSSCGGL  
 VTPDGTLLKLLIWFSTLGYTTTLLIVAFVTSALRKEELAGDELHLSVPSLPAAP  
 PSMTLAGVIDRAKOPTGTVMVAVVEGVKRVAVAGDGVVAVAGTVAATITCRHHS  
 FMPLCVG"  
 join(45402..45435,46888..46954,46962..47139)  
 /notes="hypothetical protein"  
 /codon\_start=1  
 /protein\_id="BAA83562.1"  
 /db\_xref="GI:5803252"

/translation="MHPVSGGVAVRPRQSGYPPVRAPASTLLPLRLALFSTATVP  
 APKGVNPARPPAGLLPLPFAFLCPRHHRTARPSPATFPHPTROP"  
 join(48362..48446,48739..48863,48944..49210,49363..49372,  
 50072..50152,50236..50364,50474..50571,50645..50905,  
 51205..51338,51584..51770)  
 /notes="ESTs AU068856(C50756),AU077979(C61370) correspond  
 to a region of the predicted gene: similar to serine  
 protease (AF097709)"  
 /codon\_start=1  
 /protein\_id="BAA83563.1"  
 /db\_xref="GI:5803253"  
 /translation="MMAKARGGDLPRSFHKNSRAFKNEWISGDKILFSCSGTAEHQ  
 FTKLTATLVRNALNATTKHDDLKIOVRLDGTLYDGMAYDLNDNFVVEVTSV  
 RVQVGPFOALESPLHGVLAAGHDTGEIMVETVELAGDSKVSDEDDRLHCKISKP  
 WEVHGDFLNOEQDLDSMGYPKLPSSMLGAGMILYNFEDPDGDIYEGEWRKFSRA  
 SILNRNVALASNGFFIEMSGSKMILTASLRDSDGDKNDENUKIKVFLNNQCKE  
 GLEHNCNHYNIALSVRYALRPLNTSFDCKSSRVAVGRCFNGLMATSGRLKILAS  
 TGTLCQFLACSTCKITKAGISPLVNLGNYGMNFYDTRIGTFPLLWEEICKILAS  
 FETKWSVPKRWCRPEADESDDDKLAFDDTGQLQYSYILGRKVKLLRLTIPISVPI  
 VEAQSTDG"  
 join(55308..56114,56200..56345,56442..56822)  
 /notes="ESTs C97429(C60159),D22427(C11106),  
 AU078031(E31854),D15683(C1084) correspond to a region of  
 the predicted gene: hypothetical protein"  
 /codon\_start=1  
 /protein\_id="BAA83564.1"  
 /db\_xref="GI:5803254"  
 /translation="MRAARAACVSRFSLSWRCHPNLNFSEARGLNNAAGKEELA  
 GLFYKVDHILKRGHIGVKYKTKIYSDYSGKSSYLNWLOIGVYKPGIEBLISLT  
 QFQANFPFCSLLNSGSDSIQYHLNSCSFHTVTLGLRLTLRYLCCLVITNEL  
 SCFLSHALEQLEIRYCDRIVCLVPCQLRIILKVFCDNLKLIENAPNVSTFA  
 FQGDTKLQKGETLQWKLSCVRSVYVHARALPSIMENLKSAMKSEKAFAPKL  
 CSKFLCHLSIALIGYPAYDLSIASYIHAAPSLETFLYTMORYVQNVISFAHPA  
 DARSREGHSLKSVRYTSFISVSLVELTCHILESTSLCCLDASQTSFCDTP  
 GSKICPPDLRDIIMEGRGVLAIRYIQRPVSTVKLVNLEPCSCHCTEL"  
 complement(join(58107..58328,58503..58564,58688..58751))  
 /notes="ESTs C98096(C0688),C98097(C0688) correspond to a  
 region of the predicted gene: similar to OsenOD93a gene  
 for early nodulin (AB018375)"  
 /codon\_start=1  
 /protein\_id="BAA83565.1"  
 /db\_xref="GI:5803255"  
 /translation="MATVTRAHLQRLALAKRCSREANIAGVKAATVATIASAVPTLA  
 SVRMLPWAKANINPTGQALIICTAAGMAYFVAADKKILSLARRHSFENAPHLKNTSF  
 QGTGRPHPAFFRP"  
 complement(join(64318..64539,64634..64695,64782..64845))  
 /notes="ESTs C98280(C1391),D15843(C1391) correspond to a  
 region of the predicted gene: similar to OsenOD93a gene  
 for early nodulin (AB018375)"  
 /codon\_start=1  
 /protein\_id="BAA83566.1"  
 /db\_xref="GI:5803256"  
 /translation="MSTVTRAYLDOKLALAKRCSRENTLAGAKAAVATVASAVPTLA  
 SVRMLPWAKANINPTGQALIICTAAGMAYFVAADKKILSLARRHSFENAPHLKNTSF  
 QGTGRPHPAFFRP"  
 complement(join(67321..67542,67657..67718,67829..67895))  
 /notes="EST C97982(C0324) corresponds to a region of the  
 predicted gene: similar to OsenOD93a gene for early  
 nodulin (AB018375)"  
 /codon\_start=1  
 /protein\_id="BAA83567.1"

alignment\_scores:  
 Quality: 50.00 Length: 14  
 Ratio: 3.846 Gaps: 0  
 Percent Similarity: 92.857 Percent Identity: 71.429  
 alignment\_block:  
 US-08-653-294-22 x AP000399/rev ..  
 Align seg 1/1 to reverse of: AP000399 from: 1 to: 154180  
 2 TyrLysLeuValIleLysAsnAsnIleArgIleValVal 15



```
repeat_region 8386..8487 /rpt_family="Alu"
repeat_region 8505..8621 /rpt_family="L2"
repeat_region 8631..8906 /rpt_family="L2"
repeat_region 10122..10427 /rpt_family="Alu"
repeat_region 13461..13532 /rpt_family="MIR"
repeat_region 14028..14317 /rpt_family="Alu"
repeat_region 15788..15873 /rpt_family="MIR"
repeat_region 15892..15990 /rpt_family="AT_rich"
repeat_region 17036..17079 /rpt_family="L2"
repeat_region 17162..17475 /rpt_family="Alu"
repeat_region 17585..17662 /rpt_family="MIR"
repeat_region 17668..18021 /rpt_family="MIR"
repeat_region 18022..18889 /rpt_family="MIR"
repeat_region 18929..19021 /rpt_family="L2"
repeat_region 19026..19310 /rpt_family="L2"
repeat_region 20011..20303 /rpt_family="Alu"
repeat_region 20809..20929 /rpt_family="L2"
repeat_region 21034..21303 /rpt_family="L2"
repeat_region 21630..22539 /rpt_family="L1"
repeat_region 22715..22753 /rpt_family="AT_rich"
repeat_region 23817..23968 /rpt_family="Other"
repeat_region 23969..24276 /rpt_family="Alu"
repeat_region 24277..24704 /rpt_family="Other"
repeat_region 24857..24990 /rpt_family="Other"
repeat_region 25266..25354 /rpt_family="Other"
repeat_region 25427..25458 /rpt_family="CA)n"
repeat_region 25460..25499 /rpt_family="TA)n"
repeat_region 25820..25934 /rpt_family="MIR"
repeat_region 26011..26258 /rpt_family="Other"
repeat_region 26531..26718 /rpt_family="MIR1_type"
repeat_region 27988..30411 /rpt_family="L1"
repeat_region 30412..30705 /rpt_family="Alu"
repeat_region 30706..30971 /rpt_family="L1"
repeat_region 31074..31103 /rpt_family="CA)n"
repeat_region 31196..31376 /rpt_family="TATATG)n"
repeat_region 31291..31398 /rpt_family="MIR"
repeat_region 31809..32115
```

```
repeat_region 32243..32377 /rpt_family="MaLR"
repeat_region 32380..32584 /rpt_family="L1"
repeat_region 32585..32889 /rpt_family="Alu"
repeat_region 32937..33044 /rpt_family="L1"
repeat_region 34485..34775 /rpt_family="Alu"
repeat_region 36842..36951 /rpt_family="L2"
repeat_region 37044..37346 /rpt_family="Alu"
repeat_region 37482..37489 /rpt_family="AT_rich"
repeat_region 38366..38563 /rpt_family="Alu"
repeat_region 38564..38662 /rpt_family="Alu"
repeat_region 39088..39124 /rpt_family="AT_rich"
repeat_region 39130..39223 /rpt_family="Alu"
repeat_region 39279..39584 /rpt_family="Alu"
repeat_region 39702..40055 /rpt_family="MaLR"
repeat_region 42720..43045 /rpt_family="Alu"
repeat_region 43415..43487 /rpt_family="Alu"
repeat_region 45006..45724 /rpt_family="L1"
repeat_region 45747..46161 /rpt_family="L2"
repeat_region 47437..48304 /rpt_family="L1"
repeat_region 48304..48773 /rpt_family="L1"
repeat_region 49403..49527 /rpt_family="L1"
repeat_region 49813..50107 /rpt_family="Alu"
repeat_region 50838..51171 /rpt_family="MIR1_type"
repeat_region 51307..51607 /rpt_family="Alu"

alignment_scores:
    Quality: 49.00 Length: 16
    Ratio: 3.769 Gaps: 0
    Percent Similarity: 81.250 Percent Identity: 56.250

alignment_block:
US-08-653-294-22 x AC010873/rev ..

Align seg 1/1 to reverse of: AC010873 from: 1 to: 172740

2 TyrLysLeuValIleLysIleAsnIleArgIleValIleValIysPhe 17
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
145531 TATAAGCTGATTTTAAATCAATTAATTCATTAATCAATGAAGTAT 145484

seq_name: gb_htg3:AC011386

seq_documentation_block:
LOCUS AC011386 217175 bp DNA HTG 06-OCT-1999
DEFINITION Homo sapiens chromosome 5 clone CIT978SKB_15P5, *** SEQUENCING IN
PROGRESS ***, 45 unordered pieces.
ACCESSION AC011386
VERSION AC011386.1 GI:6013555
KEYWORDS HTG; HTGS_PHASE1.
```



```

Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
Staphyliniformia; Staphylinidae; Steninae; Dianous.
REFERENCE
AUTHORS Ballard,J.W.O., Thayer,M.K., Newton,A.F. and Grismer,E.R.
TITLE Data sets, partitions, and characters: Philosphies and procedures
JOURNAL for analyzing multiple data sets
REFERENCE Unpublished
AUTHORS 2 (bases 1 to 465)
TITLE Ballard,J.W.O.
JOURNAL Direct Submission
SUBMITTED (28-AUG-1997) Zoology, The Field Museum, Roosevelt Rd at
Lake Shore Drive, Chicago, IL 60605, USA
FEATURES
Location/Qualifiers
1. .465
/organism="Dianous nitidulus"
/mitochondrion
/db_xref="taxon:66332"
<1. .>465
/gene="cytb"
<1. .>465
/gene="cytb"
/codon_start=1
/transl_table=5
/product="cytochrome b"
/protein_id="AAD01718.1"
/db_xref="GI:4103229"
/translation="MGQSFVGCATVITLLSAIPVLGEMIVKWLWGGFAVDNATLTFR
FTLHPFIVFLVFLVILHLSHSGSNPINSINKIPFHYEFSEKDFIGMILM
MLTLTLINPLMGDPDNFIPANPLIPIHQPEWFLFAILRSIPNKLIG"
BASE COUNT 156 a 68 c 49 g 192 t
ORIGIN

alignment_scores:
Quality: 48.00 Length: 12
Ratio: 4.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 58.333

alignment_block:
US-08-653-294-22 x AF021084/rev ..
Align seg 1/1 to reverse of: AF021084 from: 1 to: 465

5 ValileLysleAsnAsnilleArgileValVallys 16
:::|||||:|||||:|||||:|||||:|||||:
335 TTAATTAAGCTTAATATGTAAGAAATTATTATTAAG 300

seq_name: gb_inl:MSQNCATR

seq_documentation_block:
LOCUS MSQNCATR 15455 bp DNA INV 02-APR-1996
DEFINITION Anopheles quadrimaculatus NADH dehydrogenase subunits (1-4, 4L,
5-6); cytochrome oxidase subunits (1-3); adenosine triphosphatase
subunits (6,8); cytochrome b; transfer RNA; ribosomal RNA (large
and small subunits).
ACCESSION L04272
VERSION L04272.1 GI:342501
KEYWORDS NADH dehydrogenase subunit 1; NADH dehydrogenase subunit 2; NADH
dehydrogenase subunit 3; NADH dehydrogenase subunit 4; NADH
dehydrogenase subunit 4L; NADH dehydrogenase subunit 5; NADH
dehydrogenase subunit 6; adenosine triphosphatase subunit 6;
adenosine triphosphatase subunit 8; cytochrome b; cytochrome
oxidase subunit 1; cytochrome oxidase subunit 2; cytochrome
oxidase subunit 3; ribosomal RNA large subunit; ribosomal RNA small
subunit; transfer RNA-Ala; transfer RNA-Arg; transfer RNA-Asn;
transfer RNA-Asp; transfer RNA-Cys; transfer RNA-Gln; transfer
RNA-Glu; transfer RNA-Gly; transfer RNA-His; transfer RNA-Ile;
transfer RNA-Lys; transfer RNA-Met; transfer RNA-Phe; transfer
RNA-Pro; transfer RNA-Ser; transfer RNA-Thr; transfer RNA-Trp;
transfer RNA-Val.
SOURCE Mitochondrion Anopheles quadrimaculatus (strain Orlando) DNA.
ORGANISM Mitochondrion Anopheles quadrimaculatus A

```

```

Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Nematocera;
Culicoides; Culicidae; Anopheles.
REFERENCE
AUTHORS Cockburn,A.F., Mitchell,S.E. and Seawright,J.A.
TITLE Cloning of the mitochondrial genome of Anopheles quadrimaculatus
JOURNAL Arch. Insect Biochem. Physiol. 14 (1), 31-36 (1990)
MEDLINE 92190510
FEATURES
Location/Qualifiers
1. .15455
/organism="Anopheles quadrimaculatus A"
/mitochondrion
/db_xref="Orlando"
/db_xref="taxon:42377"
/note="Sequence determined in part by overlap with
internal peptide fragments."
1. .64
/note="codon recognized: AUG; putative"
/product="tRNA-Ile"
/anticodon="(pos:31. .33,aa:Ile)"
complement(66. .134)
/note="codon recognized: AAG; putative"
/product="tRNA-Gln"
/anticodon="(pos:96. .98,aa:Gln)"
134. .202
/note="codon recognized: AUG; putative"
/product="tRNA-Met"
/anticodon="(pos:164. .166,aa:Met)"
203. .1228
/note="putative"
/codon_start=1
/transl_table=5
/product="NADH dehydrogenase subunit 2"
/protein_id="AAA93540.1"
/db_xref="GI:1217659"
/translation="MKKISNNIFITLITGLTVLTISNSWLGWAGMGLEINLLSPIDLM
NDKKKLLTSESSLYKLTQAFSSILLFAIMLFFNNFLMNNYNEILLSTLL
LKGAAPFFHFWFEVMEGLSWINGLFTMQKIAPLMLISYFNFMFISIIISMLI
GSLGGLNOTSIRKLMFASINSIHGLWMLLMMNMLWMTYFLIYSLLSISIVLMNF
KLFYNOIFNISLMPNPIKLLIFLNLGLGLPPFLGFLPKWLVQLTNNGOLFILT
ISVCLTLITLFFVLRLSYIFMLNYQKNMKNNNKMSINLLNLFISISGLIMI
FMVYMIN"
1227. .1295
/note="codon recognized: UGA; putative"
/product="tRNA-Trp"
/anticodon="(pos:1257. .1259,aa:Trp)"
complement(1295. .1358)
/note="codon recognized: AAU; putative"
/product="tRNA-Cys"
/anticodon="(pos:1324. .1326,aa:Cys)"
complement(1359. .1424)
/note="codon recognized: UUA; putative"
/product="tRNA-Tyr"
/anticodon="(pos:1390. .1392,aa:Tyr)"
1423. .2964
/gene="COXI"
1423. .2964
/partial
/gene="COXI"
/note="putative"
/codon_start=1
/transl_table=5
/product="cytochrome oxidase subunit 1"
/protein_id="AAA93541.1"
/db_xref="GI:471964"
/translation="SRQWLESTNKKDITGLYFIFFGAWAGMVGTSLSILIRAEGLHPCA
FIGDQYINIVTAAHAFIMLFFVMPMIGCGFNWVPLMLGAPDMAPFPNNNSFWN
LPSLTLLIISSSVENGAGTGWTVYPLSSGIAHAGASVDLAIFSLHLAGISSILGAV
NFTITVINRAGPITLDRMPYVSVITAVLLLSLPVLGAIATMLLTDRLNLTSPF
DPAGGEPNLYQHLFWFFGLPGFMISHIITQESGKKTFFGLMGIYAML
AIGLGFIVWAHMTVGMVDTRAVFTSATMIIAVPTGKIFSLATMHTGTLTYSF
AMLWAFGEVLEFTVGGTLGVVLANSDDIVLHDTYVVAHFVYVLSNGAVFAIMAGFI
HWTPLLTGLTNPNWNLQFAMFVGVNLIFFPQHFGLGLAGMPRRISDFPDSYLANNI

```



McEwan, P., McGurk, A., McKernan, K., McLaughlin, J., Meldrim, J.,  
Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,  
Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P.,  
Stange-Rothmann, N., Stojanovic, N., Subramanian, A., Talamas, J.,  
Tesfaye, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X.,  
Wyman, D., Ye, W. J., Zimmer, A. and Zody, N.  
Direct Submission  
Submitted (27-OCT-1999) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
All repeats were identified using RepeatMasker: Smit, A.F.A. &  
Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>.

\* NOTE: This record contains 70 individual  
\* sequencing reads that have not been assembled into  
\* contigs. Runs of N are used to separate the reads  
\* and the order in which they appear is completely  
\* arbitrary. Low-pass sequence sampling is useful for  
\* identifying clones that may be gene-rich and allows  
\* overlap relationships among clones to be deduced.  
\* However, it should not be assumed that this clone  
\* will be sequenced to completion. In the event that  
\* the record is updated, the accession number will  
\* be preserved.

\* 1 674: contig of 674 bp in length  
\* gap of unknown length  
\* 675 1478: contig of 804 bp in length  
\* gap of unknown length  
\* 1479 2300: contig of 822 bp in length  
\* gap of unknown length  
\* 2301 3108: contig of 806 bp in length  
\* gap of unknown length  
\* 3107 3902: contig of 796 bp in length  
\* gap of unknown length  
\* 3903 4686: contig of 784 bp in length  
\* gap of unknown length  
\* 4687 5477: contig of 791 bp in length  
\* gap of unknown length  
\* 5478 6274: contig of 797 bp in length  
\* gap of unknown length  
\* 6275 7084: contig of 810 bp in length  
\* gap of unknown length  
\* 7085 7886: contig of 802 bp in length  
\* gap of unknown length  
\* 7887 8686: contig of 800 bp in length  
\* gap of unknown length  
\* 8687 9487: contig of 801 bp in length  
\* gap of unknown length  
\* 9488 10294: contig of 807 bp in length  
\* gap of unknown length  
\* 10295 11086: contig of 792 bp in length  
\* gap of unknown length  
\* 11087 11895: contig of 809 bp in length  
\* gap of unknown length  
\* 11896 12698: contig of 803 bp in length  
\* gap of unknown length  
\* 12699 13526: contig of 828 bp in length  
\* gap of unknown length  
\* 13527 14341: contig of 815 bp in length  
\* gap of unknown length  
\* 14342 15154: contig of 813 bp in length  
\* gap of unknown length  
\* 15155 15982: contig of 828 bp in length  
\* gap of unknown length  
\* 15983 16783: contig of 801 bp in length  
\* gap of unknown length  
\* 16784 17585: contig of 802 bp in length  
\* gap of unknown length  
\* 17586 18393: contig of 808 bp in length  
\* gap of unknown length  
\* 18394 19188: contig of 795 bp in length  
\* gap of unknown length  
\* 19189 19999: contig of 811 bp in length  
\* gap of unknown length

\* 20000 20791: contig of 792 bp in length  
\* gap of unknown length  
\* 20792 21613: contig of 822 bp in length  
\* gap of unknown length  
\* 21614 22393: contig of 780 bp in length  
\* gap of unknown length  
\* 22394 23180: contig of 787 bp in length  
\* gap of unknown length  
\* 23181 23990: contig of 810 bp in length  
\* gap of unknown length  
\* 23991 24799: contig of 809 bp in length  
\* gap of unknown length  
\* 24800 25594: contig of 795 bp in length  
\* gap of unknown length  
\* 25595 26419: contig of 825 bp in length  
\* gap of unknown length  
\* 26420 27237: contig of 818 bp in length  
\* gap of unknown length  
\* 27238 28037: contig of 800 bp in length  
\* gap of unknown length  
\* 28038 28831: contig of 794 bp in length  
\* gap of unknown length  
\* 28832 29615: contig of 784 bp in length  
\* gap of unknown length  
\* 29616 30382: contig of 767 bp in length  
\* gap of unknown length  
\* 30383 31174: contig of 792 bp in length  
\* gap of unknown length  
\* 31175 31962: contig of 788 bp in length  
\* gap of unknown length  
\* 31963 32776: contig of 814 bp in length  
\* gap of unknown length  
\* 32777 33583: contig of 807 bp in length  
\* gap of unknown length  
\* 33584 34386: contig of 803 bp in length  
\* gap of unknown length  
\* 34387 35195: contig of 809 bp in length  
\* gap of unknown length  
\* 35196 35990: contig of 795 bp in length  
\* gap of unknown length  
\* 35991 36798: contig of 808 bp in length  
\* gap of unknown length  
\* 36799 37602: contig of 804 bp in length  
\* gap of unknown length  
\* 37603 38405: contig of 803 bp in length  
\* gap of unknown length  
\* 38406 39204: contig of 799 bp in length  
\* gap of unknown length  
\* 39205 39968: contig of 784 bp in length  
\* gap of unknown length  
\* 39969 40772: contig of 804 bp in length  
\* gap of unknown length  
\* 40773 41593: contig of 821 bp in length  
\* gap of unknown length  
\* 41594 42400: contig of 807 bp in length  
\* gap of unknown length  
\* 42401 43239: contig of 839 bp in length  
\* gap of unknown length  
\* 43240 44023: contig of 784 bp in length  
\* gap of unknown length  
\* 44024 44804: contig of 781 bp in length  
\* gap of unknown length  
\* 44805 45597: contig of 793 bp in length  
\* gap of unknown length  
\* 45598 46389: contig of 792 bp in length  
\* gap of unknown length  
\* 46390 47177: contig of 788 bp in length  
\* gap of unknown length  
\* 47178 48010: contig of 833 bp in length  
\* gap of unknown length  
\* 48011 48800: contig of 790 bp in length  
\* gap of unknown length  
\* 48801 49616: contig of 816 bp in length  
\* gap of unknown length



```

* 49617 50417: contig of 801 bp in length
* 50418 51216: contig of 799 bp in length
* 51217 52016: contig of 800 bp in length
* 52017 52811: contig of 795 bp in length
* 52812 53614: contig of 803 bp in length
* 53615 54397: contig of 783 bp in length
* 54398 55198: contig of 801 bp in length
* 55199 56025: contig of 827 bp in length.
FEATURES
    source
    1..56025
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /chromosome="18"
    /map="18"
    /clone="408.H.20"
    /clone_lib="RC1-11 Human Male BAC"
BASE COUNT 17156 a 11260 c 10486 g 16692 t 431 others
ORIGIN

```

```

alignment_scores:
    Quality: 48.00      Length: 16
    Ratio: 3.692       Gaps: 0
    Percent Similarity: 81.250   Percent Identity: 62.500
alignment_block:
    US-08-653-294-22 x AC012417/rev
    Align seg 1/1 to reverse of: AC012417 from: 1 to: 56025

```

```

2 TyTlysLeuValIleAsnAsnIleArgIleValValIysPhe 17
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
41404 TATAAATTAATATTATATAATATATACACATTAAAGTTTCTTTT 41357

```

```
seq_name: em_in:AC004375
```

```
seq_documentation_block:
```

```

ID AC004375 standard; DNA; INV; 87325 BP.
XX AC AC004375; AC003631; AC003632; AC003633; AC003634; AC003635;
SV AC004375.1

```

```

XX 13-MAR-1998 (Rel. 55, Created)
DT 08-JUL-1998 (Rel. 56, Last updated, Version 9)
XX Drosophila melanogaster DNA sequence (P1 DS02190 (D82)), complete
DE sequence.
XX HTG.

```

```

XX OS Drosophila melanogaster (fruit fly)
XX Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota;
XX OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea;
XX Drosophilidae; Drosophila.
XX [1]
RN 1-87325
RA Celniker S.E., George R.A., Galle R.F., Hoskins R.A., Svirskas R.R.,
RA Harris N.L., Agbayani A., Arcaina T.T., Baxter E., Blazej R.G., Chavez C.,
RA Chew M., Doyle C.M., Farfan D.E., Flanagan J., Houston K.A., Hummasti S.R.,
RA Karra K., Kearney L., Kim S.H., Lee B., Lomotan M.A., Mak J., Mazda P.,
RA Mok M.S., Moshrefi A.R., Moshrefi M., Nixon K., Pacleb J.M., Park S.,
RA Pfeiffer B., Punch E., Snir E., Twomey B., Wan K.H., Whitelaw K.R., Yee A.,
RA Zhang R., Zieran L.L., Kimmel B.E.

```

```

RT "Sequencing of Drosophila chromosome, region 23C1-23C5";
RL Unpublished.
XX [2]
RN 1-87325
RA Celniker S.E., George R.A., Galle R., Svirskas R.R., Hoskins R.A.,
RA Agbayani A., Arcaina T.T., Baxter E., Blazej R.G., Chavez C., Chew M.,
RA Doyle C.M., Farfan D.E., Flanagan J., Houston K.A., Hummasti S.R.,
RA Karra K., Kearney L., Kim S.H., Lee B., Lomotan M.A., Mak J., Mazda P.,
RA Mok M.S., Moshrefi A.R., Moshrefi M., Nixon K., Pacleb J.M., Park S.,
RA Pfeiffer B., Punch E., Snir E., Twomey B., Wan K.H., Whitelaw K.R., Yee A.,
RA Zhang R., Zieran L.L., Kimmel B.E.;
RT Submitted (07-MAR-1998) to the EMBL/GenBank/DBJ databases.
RL Drosophila Genome Center, Lawrence Berkeley Laboratory, MS 64-121,
RL Berkeley, CA 94720, USA
XX Sequence submitted by:
CC Berkeley Drosophila Genome Project
CC Lawrence Berkeley National Laboratory, MS 64-121
CC Berkeley, CA 94720
CC For further information about this sequence, including its location
CC and relationship to other sequences, please visit our sequence
CC archive Web site (http://fruitfly.berkeley.edu/sequence/) or send
CC email to drosophila@hgc.lbl.gov.
CC Library location: 78-23.
XX Key Location/Qualifiers
FH source 1..87325
FT /chromosome="2L"
FT /db_xref="taxon:7227"
FT /organism="Drosophila melanogaster"
FT /map="23C1-23C5"
FT /clone="P1 DS02190 (D82)"
XX SQ Sequence 87325 BP; 23205 A; 19849 C; 20663 G; 23608 T; 0 other;

```

```
alignment_scores:
```

```

    Quality: 48.00      Length: 12
    Ratio: 4.000       Gaps: 0
    Percent Similarity: 100.000   Percent Identity: 66.667

```

```
alignment_block:
```

```

US-08-653-294-22 x AC004375/rev
    Align seg 1/1 to reverse of: AC004375 from: 1 to: 87325

```

```
6 IleLysIleAsnAsnIleArgIleValValIysPhe 17
```

```
|||||:|||||:|||||:|||||:|||||:|||||:|||||
76494 ATTAACCTAAACAACGTAGAAATCATAGTTAAATTT 76459

```

```
seq_name: gb_in2:AC007765
```

```
seq_documentation_block:
```

```

LOCUS AC007765 163403 bp DNA INV 09-JUN-1999
DEFINITION Drosophila melanogaster, chromosome 2L, region 23C1-23C5, P1 clones
DS02190 and DS0906, complete sequence.
ACCESSION AC007765 AC004375 AC003631 AC003632 AC003633 AC003634 AC003635
AC004154 AC002028 AC002976 AC002977 AC002978 AC002979 AC002980
AC002032 AC002033 AC002562 AC002034 AC002035
VERSION AC007765.1 GI:5030435
KEYWORDS HTG.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster

```

```

Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae;
1 (bases 1 to 163403)
REFERENCE
    Celniker S.E., Agbayani A., Arcaina T.T., Baxter E., Blazej R.G.,
    Bellenhof, C., Champe, M., Chavez, C., Chew, M., Ciesiolka, L.,
    Doyle, C.M., Farfan D.E., Galle, R., George, R.A., Harris, N.L.,

```



```
gene
CDS
/gene="ORF10"
1805..2026
/gene="ORF18"
1805..2026
/gene="ORF18"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="hypothetical protein"
/protein_id="CAA72059.1"
/db_xref="GI:2239067"
/db_xref="SPTREMBL:O32310"
/translacion="MTVADIKKKKANYQFTTTPSEMELIEIIVDIENDRRKTIARENN
LPFKYNTFTFLAMIEKKRIYQGEI"
2026..2322
/gene="ORF1"
2026..2322
/gene="ORF1"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="hypothetical protein"
/protein_id="CAA72054.1"
/db_xref="GI:2239068"
/db_xref="SPTREMBL:O32311"
/translacion="MKYSKTLFLDFENQVVVDVTKNWKWDCMAPDDFAEYVYII
IENARRYKFHFDEPDYMFNLSPYKVNLFNGENYILIMNPFESQLVYN"
2338..2562
/gene="ORF14"
2338..2562
/gene="ORF14"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="hypothetical protein"
/protein_id="CAA72055.1"
/db_xref="GI:2239069"
/db_xref="SPTREMBL:O32312"
/translacion="MMEDLLHKIIDKDKIDQIEIKTDLKAHRIETDDNFTKLYHL
VEEKDIEDVNLKRLFKNEATVTERIKREKH"
3210..4133
/gene="rep"
3210..4133
/gene="rep"
/feature="rep"
/codon_start=1
/transl_table=11
/evidence=experimental
/product="replication protein"
/protein_id="CAA72056.1"
/db_xref="GI:2239070"
/db_xref="SPTREMBL:O32313"
/translacion="MGHILITGSMKLLRSDYGFSDIGLLKFLSSISKERHFSVDV
AKDTSQSVSKKARYIKDGNLITFRFGGQIKFAPKLIGEEDKLYQVDPGETWYL
GSRSTQRFYDKKQAMNADLLHWTCEQLQVDDAATNFKNVILVDKLFKEPCYS
FLTYVDFKTKTGAKHMRNDSAKFWADFLDSVEKVKLGRKKKSLDDVNPVYLWE
KLLGDADFYESMSQEKDKFKWFYEQIISVLYLRALKHKEGVVEEYFKKLIHGM
RVDQKKVAVGVQLLMSERLKKIKEESSIQ"
4614..4925
/gene="ORF7"
4614..4925
/gene="ORF7"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="hypothetical protein"
/protein_id="CAA72057.1"
/db_xref="GI:2239071"
/db_xref="SPTREMBL:O32314"
/translacion="MFDFIVKLYGGIIMPVQIEGVFMGASIIITKEFGNKTSL
IDLQKNNGNETVKISSKEVELFOQISEGKLGIPFKLATVSAKNNAYINLQOIV
R"
5988..7415
gene
CDS
/gene="ORF8"
5988..7415
/gene="ORF8"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="hypothetical protein"
/protein_id="CAA72058.1"
/db_xref="GI:2239072"
/db_xref="SPTREMBL:O32315"
/translacion="MKKLAIFELAFCAIVSHITPVFADSPSFQVSGKNVTIENNGIV
NGKNLIIKVNMMKDVYSQSELMVLVNDSCFHVGSKWTRPEQKNEFPINWNGGT
YAFYFFHGVYGFARFKLGKFGKNGGGGTYYWGKAKYDDEIFEKPNEDYADY
DNKNDGVTNEKIPDKPQGGDIDESKDNKPPVDGGGSDKAVKEALEKIVNAVNE
ISKNTGETAKNTGETAKNTGETAKNTGETAKNTGETAKNTGETAKNTGETAKNTGET
DMDLSIKPVLPEDIKPKDPKOPDFEDKQEHFKEGQEKPPDKLPVAPPEVCEKGDV
CKDNOKEDMEKDKMDKEPMKPKMEKEDMKPKMTPDKPLEKEDMKPKPM
TPDKPLEKEDMKPKMTPDKPLEKENMKPKMTPDKPMDREGEMKNDNMKQDKPM
TPWEEMKPSDEGSELCHWHVNGKCG"
7060..7280
/feature="repeat is probably at protein level"
/rpt_type=TANDEM
/evidence=not_experimental
7406..8809
/gene="ORF4"
7406..8809
/gene="ORF4"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="hypothetical protein"
/protein_id="CAA72049.1"
/db_xref="GI:2239073"
/db_xref="SPTREMBL:O32316"
/translacion="MRIKFFIVPVLLGILFVGDTSAFQAGKDVLSRGVGAQYGS
YITSLDGNLQTSLSNGEVIIDIASPVRSVGFRTKVGSSNRMDVIFYTEGATLK
REYRNTDSYLOAVDIDGKVKIGLNKNSGDYFTLAENVLYEYSPSPDITDLKENI
KHSVEFYTKPKNSFNHRIYRDQVLASQKKEKFDKGLTPEHTEYKFSVSPD
GLESGIKKIKTVKEPLDPEIKQLODEIHTYVFKYKPKNNFHNKPYRDDKVI
ADNVVEEFDIKGLKSGMEYTKFVSVSSTGLENGIVKVKTDEPNLKPFPKPSNF
DYTPKDSLVNLLSQNLGVIKGYHIDQVKVNDLSLVTSRSYAVKGLQNGQTYGVQ
IKAVSKNIESDLSNVNNGIPKVPVPEVTFSGLGALVDGIIITNFWLWIPVAFSIA
IPLAFIVAFNTKKLFMR"
8971..9393
/gene="ORF2"
8971..9393
/gene="ORF2"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="hypothetical protein"
/protein_id="CAA72050.1"
/db_xref="GI:2239074"
/db_xref="SPTREMBL:O32317"
/translacion="MMKFKVRLKGGFEAVIKTKTFIDFLATPFQYLLDLFLKGIYF
IERFQVVVLLIKLIFALFOFLGAVIAGFFKTLFWQPNFDGNGVSPSASMRGFKY
ATDILLPGLLTVIPAVVTCITWLVFKKLMDLFRGGH"
10074..10436
/gene="ORF9"
10074..10436
/gene="ORF9"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="hypothetical protein"
/protein_id="CAA72051.1"
/db_xref="GI:2239075"
/db_xref="SPTREMBL:O32318"
/translacion="MLAKTRKQVQPEGHGHALIKKEGTCVVEQVVSFDSILETST
AMLPRENLRLDDNHGVYHYINLEKPARIEAENLMLRSSALKGIFDFTGKKPF
DIIGMPWIIILAIKFR"
10450..10812
/gene="ORF3"
gene
```



FEATURES  
source  
\* 77723 80958: contig of 3236 bp in length.  
Location/Qualifiers  
1. .80958  
/organism="Drosophila melanogaster"  
/db\_xref="taxon:7227"  
/chromosome="3L/73B5"  
/clone="RPCI98-11G5"  
BASE COUNT 22949 a 17495 c 17230 g 23219 t 65 others  
ORIGIN

alignment\_scores:  
Quality: 47.00 Length: 17  
Ratio: 3.615 Gaps: 0  
Percent Similarity: 76.471 Percent Identity: 47.059

alignment\_block:  
US-08-653-294-22 x AC010688 ..  
Align seg 1/1 to: AC010688 from: 1 to: 80958

1 SerTyrLysLeuValIleLysIleAsnIleArgIleValValLys 17  
|||||:||||| ||| |||||||:|||||: |||||  
32302 TCCTACAGCCTTTATACATATACAACTTACAGTTATTACAAATT 32351

17 e 17

32352 T 32352

seq\_name: gb\_pl1:AB006698

seq\_documentation\_block:  
LOCUS AB006698 84510 bp DNA PLN 20-NOV-1999  
DEFINITION Arabidopsis thaliana genomic DNA, chromosome 5, p1 clone: MCL19,  
complete sequence.  
ACCESSION AB006698  
VERSION AB006698.1 GI:2351063  
KEYWORDS HTG.  
SOURCE Arabidopsis thaliana (strain:Columbia) DNA, clone\_lib:Mitsui P1  
clone:MCL19.  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core  
eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;  
Arabidopsi.

REFERENCE  
1 (sites)  
Kotani,H., Nakamura,Y., Sato,S., Kaneko,T., Asamizu,E., Miyajima,N.  
and Tabata,S.  
TITLE Structural analysis of Arabidopsis thaliana chromosome 5. II.  
Sequence features of the regions of 1,044,062 bp covered by  
thirteen physically assigned p1 clones  
DNA Res. 4 (4), 291-300 (1997)

JOURNAL  
MEDLINE 98069011  
REFERENCE 2 (bases 1 to 84510)  
AUTHORS Nakamura,Y.  
TITLE Direct Submission  
JOURNAL Submitted (22-AUG-1997) to the DDBJ/EMBL/Genbank databases.  
Yasukazu Nakamura, Kazusa DNA Research Institute, Laboratory of  
Gene Structure 2; 1532-3, Yana, Kisarazu, Chiba 292, Japan  
(E-mail:ynakamura@kazusa.or.jp, Tel:++81-438-52-3935,  
Fax:++81-438-52-3934)

FEATURES  
source  
1. .84510  
Location/Qualifiers  
/organism="Arabidopsis thaliana"  
/strain="Columbia"  
/db\_xref="taxon:3702"  
/chromosome="5"  
/clone="MCL19"  
/clone\_lib="Mitsui P1"  
BASE COUNT 28311 a 14338 c 14358 g 27503 t  
ORIGIN

alignment\_scores:  
Quality: 47.00 Length: 13  
Ratio: 3.917 Gaps: 0  
Percent Similarity: 92.308 Percent Identity: 69.231  
alignment\_block:  
US-08-653-294-22 x AB006698/rev ..  
Align seg 1/1 to reverse of: AB006698 from: 1 to: 84510

4 LeuValIleLysIleAsnIleArgIleValValLys 16

|||||:|||||:|||||:|||||:|||||:|||||  
64930 TTAAAGTAAATAATATACATAAAGCTGGTAGTAAG 64892

seq\_name: gb\_htg4:AC008308

seq\_documentation\_block:

LOCUS AC008308 151373 bp DNA HTG 20-OCT-1999  
DEFINITION Drosophila melanogaster chromosome 3 clone BACR10M16 (D743) RPCI-98  
10.M.16 map 93C-93D strain y; cn bw sp, \*\*\* SEQUENCING IN PROGRESS  
\*\*\*, 186 unordered pieces.

ACCESSION AC008308  
VERSION AC008308.3 GI:6087905  
KEYWORDS HTG; HTGS\_PHASE1.  
SOURCE fruit fly.

ORGANISM

Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE

AUTHORS  
1 (bases 1 to 151373)  
Celisner,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,  
Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,  
Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,  
Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L.,  
Kim,E., Lee,B., Lewis,S., Li,P., Lomotan,M.A., Mazda,P.,  
Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S.,  
Pfeiffer,B., Poon,L., Sequeira,A., Sethi,H., Snir,E.,  
Svirskas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and  
Rubin,G.M.

Sequencing of Drosophila melanogaster

TITLE

JOURNAL

REFERENCE

AUTHORS  
2 (bases 1 to 151373)  
Celisner,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,  
Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,  
Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,  
Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L.,  
Kim,E., Lee,B., Lewis,S., Li,P., Lomotan,M.A., Mazda,P.,  
Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S.,  
Pfeiffer,B., Poon,L., Sequeira,A., Sethi,H., Snir,E.,  
Svirskas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and  
Rubin,G.M.

Direct Submission

TITLE

JOURNAL

COMMENT

Submitted (02-AUG-1999) Drosophila Genome Center, Lawrence Berkeley  
Laboratory, MS 64-121, Berkeley, CA 94720, USA  
On Oct 20, 1999 this sequence version replaced gi:6067129.  
For further information about this sequence, including its location  
and relationship to other sequences, please visit our sequence  
archive web site (<http://www.fruitfly.org/sequence/>) or send email  
to [bagp@fruitfly.berkeley.edu](mailto:bagp@fruitfly.berkeley.edu). All contigs in this submission meet  
the following cutoffs: length >= 200 bases.  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 186 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence.  
\* as soon as it is available and the accession number will  
\* be preserved.

\* 1 270: contig of 270 bp in length  
\* 271 350: gap of unknown length  
\* 351 1076: contig of 726 bp in length  
\* 1077 1156: gap of unknown length  
\* 1157 1787: contig of 631 bp in length

1788	1867: gap of unknown length	31978	32578: contig of 601 bp in length
1868	2168: contig of 301 bp in length	32579	32658: gap of unknown length
2169	2248: gap of unknown length	33384	3384: contig of 726 bp in length
2249	2801: contig of 553 bp in length	33464	3464: gap of unknown length
2802	2881: gap of unknown length	33465	35293: contig of 1829 bp in length
2882	3182: contig of 301 bp in length	33573	3573: gap of unknown length
3183	3262: gap of unknown length	35294	36220: contig of 847 bp in length
3263	3885: contig of 623 bp in length	36221	36300: gap of unknown length
3886	3965: gap of unknown length	36301	37223: contig of 1023 bp in length
3966	4450: contig of 485 bp in length	37324	37403: gap of unknown length
4451	4530: gap of unknown length	37404	38441: contig of 1038 bp in length
4531	5295: contig of 765 bp in length	38442	38521: gap of unknown length
5296	5375: gap of unknown length	38522	39603: contig of 1082 bp in length
5376	6199: contig of 824 bp in length	39604	39683: gap of unknown length
6200	6279: gap of unknown length	39684	40530: contig of 847 bp in length
6280	7185: contig of 906 bp in length	40531	40610: gap of unknown length
7186	7265: gap of unknown length	40611	41832: contig of 1222 bp in length
7266	7844: contig of 579 bp in length	41833	41912: gap of unknown length
7845	7924: gap of unknown length	41913	43390: contig of 1478 bp in length
7925	8688: contig of 764 bp in length	43391	43470: gap of unknown length
8689	8768: gap of unknown length	43471	44364: contig of 894 bp in length
8769	9397: contig of 629 bp in length	44365	44444: gap of unknown length
9398	9477: gap of unknown length	44445	45927: contig of 1483 bp in length
9478	9994: contig of 517 bp in length	45928	46007: gap of unknown length
9995	10074: gap of unknown length	46008	47580: contig of 1573 bp in length
10075	10705: contig of 631 bp in length	47581	47660: gap of unknown length
10706	10785: gap of unknown length	47661	48605: contig of 945 bp in length
10786	11807: contig of 1022 bp in length	48606	48685: gap of unknown length
11808	11887: gap of unknown length	48686	50085: contig of 1400 bp in length
11888	12654: contig of 767 bp in length	50086	50165: gap of unknown length
12655	12734: gap of unknown length	50166	51444: contig of 1279 bp in length
12735	13432: contig of 698 bp in length	51445	51524: gap of unknown length
13433	13512: gap of unknown length	51525	53165: contig of 1641 bp in length
13513	14324: contig of 812 bp in length	53166	53245: gap of unknown length
14325	14404: gap of unknown length	53246	54634: contig of 1389 bp in length
14405	15262: contig of 858 bp in length	54635	54714: gap of unknown length
15263	15342: gap of unknown length	54715	56896: contig of 2182 bp in length
15343	16038: contig of 696 bp in length	56897	56976: gap of unknown length
16039	16118: gap of unknown length	56977	58340: contig of 1364 bp in length
16119	17043: contig of 925 bp in length	58341	58420: gap of unknown length
17044	17123: gap of unknown length	58421	59885: contig of 1565 bp in length
17124	17775: contig of 652 bp in length	59886	60065: gap of unknown length
17776	17855: gap of unknown length	60066	63185: contig of 3120 bp in length
17856	18158: contig of 303 bp in length	63186	63265: gap of unknown length
18159	18238: gap of unknown length	63266	65461: contig of 2196 bp in length
18239	18767: contig of 529 bp in length	65462	65541: gap of unknown length
18768	18847: gap of unknown length	65542	67549: contig of 2008 bp in length
18848	20191: contig of 1344 bp in length	67550	67629: gap of unknown length
20192	20271: gap of unknown length	67630	70854: contig of 3225 bp in length
20272	20785: contig of 514 bp in length	70855	70934: gap of unknown length
20786	20865: gap of unknown length	70935	75464: contig of 4530 bp in length
20866	21878: contig of 1013 bp in length	75465	75544: gap of unknown length
21879	21958: gap of unknown length	75545	76231: contig of 687 bp in length
21959	22626: contig of 668 bp in length	76232	78311: gap of unknown length
22627	22706: gap of unknown length	76312	78866: contig of 555 bp in length
22707	23349: contig of 643 bp in length	76867	79546: gap of unknown length
23350	23429: gap of unknown length	76947	77554: contig of 608 bp in length
23430	24076: contig of 647 bp in length	77555	77634: gap of unknown length
24077	24156: gap of unknown length	77635	78331: contig of 597 bp in length
24157	25336: contig of 1180 bp in length	78232	78311: gap of unknown length
25337	25416: gap of unknown length	78312	78927: contig of 616 bp in length
25417	26672: contig of 1256 bp in length	78928	79007: gap of unknown length
26673	26752: gap of unknown length	79008	79649: contig of 642 bp in length
26753	27512: contig of 760 bp in length	79650	79729: gap of unknown length
27513	27592: gap of unknown length	79730	80365: contig of 636 bp in length
27593	28805: contig of 1213 bp in length	80366	80445: gap of unknown length
28806	28885: gap of unknown length	80446	80780: contig of 335 bp in length
28886	29725: contig of 840 bp in length	80781	80860: gap of unknown length
29726	29805: gap of unknown length	80861	81485: contig of 625 bp in length
29806	30552: contig of 747 bp in length	81486	81565: gap of unknown length
30553	30632: gap of unknown length	81566	81905: contig of 340 bp in length
30633	31897: contig of 1265 bp in length		
31898	31977: gap of unknown length		

alignment\_scores:



misc_feature	complement(27838..28143)	zbr97b08.s1	complement(27887..28143)	/notes="similar to human EST W31561 (NTD:g1312680)
misc_feature	complement(27838..28143)	zbr97b08.s1	complement(27887..28143)	/notes="similar to human EST W31561 (NTD:g1312680)
misc_feature	complement(27997..28144)	zbr97c08.s1	complement(27997..28144)	/notes="similar to human EST W31628 (NTD:g1312688)
misc_feature	28666..28972	yf26d05.s1	28666..28972	/notes="similar to human EST R09339 (NTD:g761262)
misc_feature	complement(29587..29853)	yf26d05.s1	complement(29587..29853)	/notes="similar to human EST R09227 (NTD:g761150)
misc_feature	30146..30314	yv71e06.s1	30146..30314	/notes="similar to human EST N78042 (NTD:g1240743)
misc_feature	complement(31170..31350)	yv71e06.s1	complement(31170..31350)	/notes="similar to human EST N58451 (NTD:g1202341)
repeat_region	31970..32006	rpt_family="L1"	31970..32006	rpt_family="L1"
repeat_region	32263..32563	rpt_family="ALU"	32263..32563	rpt_family="ALU"
repeat_region	complement(34737..34778)	rpt_family="L1"	complement(34737..34778)	rpt_family="L1"
repeat_region	36558..36709	rpt_family="ALU"	36558..36709	rpt_family="ALU"
repeat_region	complement(37630..37740)	rpt_family="PTR"	complement(37630..37740)	rpt_family="PTR"
repeat_region	39839..40125	rpt_family="ALU"	39839..40125	rpt_family="ALU"
repeat_region	40126..40157	rpt_family="L1"	40126..40157	rpt_family="L1"
repeat_region	complement(41340..41706)	rpt_family="ALU"	complement(41340..41706)	rpt_family="ALU"
repeat_region	complement(41904..41933)	rpt_family="L1"	complement(41904..41933)	rpt_family="L1"
repeat_region	42348..42409	rpt_family="ALU"	42348..42409	rpt_family="ALU"
repeat_region	complement(42942..42982)	rpt_family="L1"	complement(42942..42982)	rpt_family="L1"
repeat_region	complement(44380..44423)	rpt_family="L1"	complement(44380..44423)	rpt_family="L1"
repeat_region	44659..44954	rpt_family="ALU"	44659..44954	rpt_family="ALU"
repeat_region	45422..45448	rpt_family="L1"	45422..45448	rpt_family="L1"
repeat_region	complement(45615..45669)	rpt_family="L1"	complement(45615..45669)	rpt_family="L1"
repeat_region	complement(45672..45963)	rpt_family="ALU"	complement(45672..45963)	rpt_family="ALU"
repeat_region	complement(45964..46597)	rpt_family="L1"	complement(45964..46597)	rpt_family="L1"
repeat_region	46100..46516	rpt_family="L1"	46100..46516	rpt_family="L1"
repeat_region	48370..48785	rpt_family="L1"	48370..48785	rpt_family="L1"
repeat_region	complement(49577..50216)	rpt_family="L1"	complement(49577..50216)	rpt_family="L1"
repeat_region	complement(50540..50791)	rpt_family="THE"	complement(50540..50791)	rpt_family="THE"
repeat_region	complement(50820..50985)	rpt_family="THE"	complement(50820..50985)	rpt_family="THE"
repeat_region	complement(51124..51156)	rpt_family="L1"	complement(51124..51156)	rpt_family="L1"
repeat_region	51405..51691	rpt_family="ALU"	51405..51691	rpt_family="ALU"
repeat_region	complement(51721..51764)	rpt_family="L1"	complement(51721..51764)	rpt_family="L1"
repeat_region	complement(51765..52058)	rpt_family="ALU"	complement(51765..52058)	rpt_family="ALU"



OM of: US-08-653-294-22 to: N\_Geneseq\_36:\* out\_format : pfs

Date: Feb 8, 2000 7:30 PM

About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 CompuGen Ltd.

#### Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp  
-O=/cgn1\_1/USPTO\_spool/US08653294/runat\_04022000\_160701\_15807/app\_query.fasta.2  
-DB=N\_Geneseq\_36 -QFMT=fastap -SUFFIX=ring -GAPOP=12.000  
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPEXT=0.000 -LOOPEXT=0.000  
-GAPOP=4.500 -GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500  
-GAPOP=6.000 -GAPEXT=7.000 -XGAPOP=10.000 -YGAPEXT=0.500  
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blossum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR\_SCORE=pct  
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0  
-MAXLEN=1000000 -USER=US08653294 -NCPU=6 -ICPU=3 -NO\_XLPXY -WAIT  
-THREADS=1

#### Search information block:

Query: US-08-653-294-22  
Query length: 17  
Database: N\_Geneseq\_36:\*  
Database sequences: 311585  
Database length: 125096042  
Search time (sec): 873.190000

#### score\_list:

Sequence	Strd Orig	Zscore	EScore	Len	Documentation
N_Geneseq_36:X20248_08	45.00	86.54	3.2e+03	110000	Continuation (9 of 10) of
N_Geneseq_36:V21209_08	42.00	77.22	1.1e+04	110000	Continuation (9 of 17) of
N_Geneseq_36:X20248_00	42.00	77.22	1.1e+04	110000	Borrelia burgdorferi polyn
N_Geneseq_36:X20248_03	42.00	77.22	1.1e+04	110000	Continuation (4 of 10) of
N_Geneseq_36:X20248_04	41.50	112.19	119.79	1673	B. fragilis DNA probe BF-35.1
N_Geneseq_36:X20248_05	41.50	75.67	1.3e+04	110000	Continuation (5 of 10) of
N_Geneseq_36:X20248_06	41.50	75.19	1.4e+04	116277	Borrelia burgdorferi polynuc
N_Geneseq_36:X20249	41.00	126.74	18.52	264	Staphylococcus aureus contig SE
N_Geneseq_36:T83707	41.00	115.35	79.80	974	DNA encoding a Staphylococcus a
N_Geneseq_36:V74433	41.00	113.58	100.23	1194	Staphylococcus aureus contig S
N_Geneseq_36:X12992	41.00	97.51	78.67	7528	Enterococcus faecalis genome c
N_Geneseq_36:X20535	41.00	88.49	2.5e+03	21170	Polynucleotide sequence from
N_Geneseq_36:V21209_05	41.00	74.12	1.6e+04	110000	Continuation (6 of 17) of
N_Geneseq_36:V21209_07	41.00	74.12	1.6e+04	110000	Continuation (8 of 17) of
N_Geneseq_36:V90892	40.00	110.44	149.81	1198	Nucleotide sequence of cluster
N_Geneseq_36:V52021	40.00	108.62	189.36	1477	Helicobacter polypeptide GHPO
N_Geneseq_36:V74635	40.00	97.93	745.94	5030	Staphylococcus aureus contig S
N_Geneseq_36:Q67860	40.00	94.69	1.1e+03	7291	ALVAC subgenomic fragment cont
N_Geneseq_36:T69982	40.00	94.62	1.1e+03	7350	Canarypox virus DNA containing
N_Geneseq_36:Q22770	40.00	94.62	1.1e+03	7351	Canarypox virus BgIII fragment
N_Geneseq_36:T47529	40.00	94.62	1.1e+03	7351	Canarypox virus DNA containing
N_Geneseq_36:X12024	40.00	93.54	1.3e+03	8318	Borrelia burgdorferi polynucle
N_Geneseq_36:X12942	40.00	93.36	1.3e+03	8486	Enterococcus faecalis genome c
N_Geneseq_36:X13193	40.00	92.30	1.5e+03	9589	Enterococcus faecalis genome c
N_Geneseq_36:X20248_09	40.00	91.33	1.7e+03	10715	Continuation (10 of 10) of
N_Geneseq_36:V21210	40.00	76.54	1.1e+04	58407	Methanococcus jannaschii larg
N_Geneseq_36:T58840_1	40.00	71.01	2.3e+04	110000	Continuation (2 of 6) of T5
N_Geneseq_36:X20248_07	40.00	71.01	2.3e+04	110000	Continuation (8 of 10) of
N_Geneseq_36:X20248_08	40.00	71.01	2.3e+04	110000	Continuation (9 of 10) of
N_Geneseq_36:X20865	39.00	115.65	76.83	462	Polynucleotide sequence from th
N_Geneseq_36:Q20395	39.00	114.14	93.18	549	Borrelia burgdorferi polynucle
N_Geneseq_36:X03870	39.00	110.97	140.02	790	H. influenzae DNA fragment conta
N_Geneseq_36:N80227	39.00	110.33	151.97	850	Sequence of Haemophilus influen
N_Geneseq_36:T41616	39.00	99.45	613.78	2960	Adzuki bean endo-xyloglucan tr
N_Geneseq_36:T59908	39.00	96.81	860.12	4002	Yeast transcription regulatory
N_Geneseq_36:V71729	39.00	94.22	1.2e+03	5391	Upstream sequence of the ACC s
N_Geneseq_36:V71738	39.00	94.19	1.2e+03	5407	Upstream sequence of the ACC s
N_Geneseq_36:T58840_4	39.00	67.91	3.4e+04	110000	Continuation (5 of 6) of T5
N_Geneseq_36:V21209_06	39.00	67.91	3.4e+04	110000	Methanococcus jannaschii c
N_Geneseq_36:T13635	39.00	66.19	4.2e+04	13894	ACNPV genomic DNA clone 6. A
N_Geneseq_36:Q11657	38.50	94.31	1.2e+03	4462	3' Terminal of TSWV L RNA. DNA

N\_Geneseq\_36:V76315 - 38.00 124.91 23.44 112 ! Staphylococcus aureus contig  
N\_Geneseq\_36:Q05757 - 38.00 116.34 70.31 299 ! Clone cBp408 with homology t  
N\_Geneseq\_36:Q05756 - 38.00 114.62 87.62 364 ! Clone cBp405 with homology t  
N\_Geneseq\_36:Q05755 - 38.00 114.16 93.02 384 ! Clone cBp401 with homology t

seq\_name: N\_Geneseq\_36:X20248\_08

#### seq\_documentation\_block:

Continuation (9 of 10) of X20248 from base 800001 (Borrelia burgdorferi polynucleotid  
WP Sequence split into 10 fragments LOCUS X20248 Accession X20248  
Fragment Name Begin End  
WP X20248\_00 1 110000  
WP X20248\_01 100001 210000  
WP X20248\_02 210001 310000  
WP X20248\_03 310001 410000  
WP X20248\_04 410001 510000  
WP X20248\_05 510001 610000  
WP X20248\_06 610001 710000  
WP X20248\_07 710001 810000  
WP X20248\_08 810001 910000  
WP X20248\_09 910001 910715

#### alignment\_scores:

Quality: 45.00 Length: 16  
Ratio: 3.462 Gaps: 0  
Percent Similarity: 81.250 Percent Identity: 62.500

#### alignment\_block:

US-08-653-294-22 x X20248\_08

Align seg 1/1 to: X20248\_08 from: 1 to: 110000

1 SerTyrLysLeuValIleLysIleAsnAsnIleArgIleValVallys 16  
|||||  
94894 |GCTACAGCTAGTACTTACTTGGAATCTATATTAATTTGCAAAA 94941

seq\_name: N\_Geneseq\_36:V21209\_08

#### seq\_documentation\_block:

Continuation (9 of 17) of V21209 from base 800001 (Methanococcus jannaschii circular  
WP Sequence split into 17 fragments LOCUS V21209 Accession V21209  
Fragment Name Begin End  
WP V21209\_00 1 110000  
WP V21209\_01 100001 210000  
WP V21209\_02 210001 310000  
WP V21209\_03 310001 410000  
WP V21209\_04 410001 510000  
WP V21209\_05 510001 610000  
WP V21209\_06 610001 710000  
WP V21209\_07 710001 810000  
WP V21209\_08 810001 910000  
WP V21209\_09 910001 1010000  
WP V21209\_10 1010001 1110000  
WP V21209\_11 1110001 1210000  
WP V21209\_12 1210001 1310000  
WP V21209\_13 1310001 1410000  
WP V21209\_14 1410001 1510000  
WP V21209\_15 1510001 1610000  
WP V21209\_16 1610001 1664976

#### alignment\_scores:

Quality: 42.00 Length: 16  
Ratio: 3.500 Gaps: 0  
Percent Similarity: 75.000 Percent Identity: 56.250

#### alignment\_block:

US-08-653-294-22 x V21209\_08

Align seg 1/1 to: V21209\_08 from: 1 to: 110000

2 TyrLysLeuValIleLysIleAsnAsnIleArgIleValVallysPhe 17

||||| ||||| ::||| ||||| ||||| ::|||  
71863 TATACCAAGTCCTCTTATATATCAAAATAGTTTTTAACCTT 71910

seq\_name: N\_Geneseq\_36:X20248\_00

seq\_documentation\_block:

WP Sequence split into 10 fragments LOCUS X20248 Accession X20248

WP Fragment Name Begin End

WP X20248\_00 1 110000

WP X20248\_01 100001 210000

WP X20248\_02 200001 310000

WP X20248\_03 300001 410000

WP X20248\_04 400001 510000

WP X20248\_05 500001 610000

WP X20248\_06 600001 710000

WP X20248\_07 700001 810000

WP X20248\_08 800001 910000

WP X20248\_09 900001 910715

ID X20248 standard; DNA; 910715 BP.

AC X20248:

DT 04-MAY-1999 (first entry)

DE Borrelia burgdorferi polynucleotide sequence #1.

KW Borrelia burgdorferi; Spirochete; Bacterium; Pathogen; Lyme disease;

KW epidemic relapsing fever; endemic relapsing fever; Lyme borreliosis;

KW infection; diagnosis; characterisation; detection; ds.

OS Borrelia burgdorferi.

PN W09858943-A1.

PD 30-DEC-1998.

PF 18-JUN-1998; U12764.

PR 03-SEP-1997; US-057483.

PR 20-JUN-1997; US-050359.

PR 22-JUL-1997; US-053344.

PR 22-JUL-1997; US-053377.

PA (HUMA-) HUMAN GENOME SCI INC.

FA (MED-) MEDIMMUNE INC.

PI Clayton R, Dougherty BA, Fraser C, Lathigra R, Smith HO,

PI White OR;

DR WPI; 99-081217/07.

PT New isolated Borrelia burgdorferi nucleic acids - used to develop

PT products for the detection, diagnosis, characterisation, prevention

PT and therapy of infections, particularly Lyme disease

PS Claim 1; Page 157-671; 1128pp; English.

CC X20248 to X20402 represent polynucleotide sequences isolated from

CC Borrelia burgdorferi (Bb). Products derived from Bb can be used for

CC the detection, diagnosis, characterisation, prevention and therapy of

CC Bb infections, e.g. Lyme disease. They can also be used for the

CC production of biosynthetic products, e.g. enzymes. Borrelia belongs

CC to a family of motile, spiral-shaped bacteria called Spirochetes.

CC Spirochetes are pathogenic in humans and Borrelia causes epidemic and

CC endemic relapsing fever, and Lyme borreliosis, more commonly known as

CC Lyme disease.

SQ Sequence 910715 BP; 327171 A; 129646 C; 130753 G; 323091 T;

alignment\_scores:

Quality: 42.00 Length: 14

Ratio: 3.818 Gaps: 0

Percent Similarity: 78.571 Percent Identity: 71.429

alignment\_block:

US-08-653-294-22 x X20248\_00/rev ..

Align seg 1/1 to reverse of: X20248\_00 from: 1 to: 110000

3 LysLeuValIleAlaGlnAsnIleArgIleValValLys 16

||||| ||||| ||||| ||||| |||||

44990 AAACCTTTTAAATAATCTTGATATGAGATATGCTGAAG 44949

seq\_name: N\_Geneseq\_36:X20248\_03

seq\_documentation\_block:

Continuation (4 of 10) of X20248 from base 300001 (Borrelia burgdorferi polynucleotide s

WP Sequence split into 10 fragments LOCUS X20248 Accession X20248

WP Fragment Name Begin End

WP X20248\_00 1 110000

WP X20248\_01 100001 210000

WP X20248\_02 200001 310000

WP X20248\_03 300001 410000

WP X20248\_04 400001 510000

WP X20248\_05 500001 610000

WP X20248\_06 600001 710000

WP X20248\_07 700001 810000

WP X20248\_08 800001 910000

WP X20248\_09 900001 910715

alignment\_scores:

Quality: 42.00 Length: 17

Ratio: 3.500 Gaps: 0

Percent Similarity: 70.588 Percent Identity: 52.941

alignment\_block:

US-08-653-294-22 x X20248\_03 ..

Align seg 1/1 to: X20248\_03 from: 1 to: 110000

1 SerTyrLysLeuValIleAlaGlnAsnIleArgIleValValLysPh 17

:::||||| ::| ||||| ||||| |||||

67263 ACATACAAGAGATCTGCAAGATTTTCAACATAAGATAAAATCAAGTT 67312

17 e 17

67313 T 67313

seq\_name: N\_Geneseq\_36:X22323

seq\_documentation\_block:

ID X22323 standard; DNA; 1673 BP.

AC X22323;

DT 19-MAY-1999 (first entry)

DE B. fragilis DNA probe BF-35.

KW Probe; diagnosis; infection; specific; blood; detection; BF-35; ds.

OS Bacterioides fragilis.

PN W09842844-A1.

PD 01-OCT-1998.

PF 23-MAR-1998; J01287.

PR 25-MAR-1997; JP-071079.

PA (FUSO ) FUSO PHARM IND LTD.

PI (FUSO ) FUSO YAKUHIN KOGYO KK.

PI Abe K, Keshi H, Matsuhisa A, Ueyama H;

DR WPI; 99-070072/06.

PT DNA probes for diagnosis of infection with Bacteroides fragilis -

PT are obtained by HindIII digestion of genomic B. fragilis DNA and

PT selection of bacterial-binding fragments

PS Claim 2; Page 25-26; 36pp; Japanese.

CC X22318-X22323 are DNA probes used for the diagnosis of infection with

CC Bacteroides fragilis. The probes are obtained by isolation of genomic

CC DNA from this bacterium, digesting it with HindIII restriction enzyme,

CC and identifying clones capable of binding to the bacterium then isolating

CC and sequencing the HindIII fragments involved. The method allows simple

CC and highly specific detection of Bacteroides fragilis in biological

CC samples such as blood, for diagnosis of Bacteroides fragilis infections.

CC This sequence represents genomic DNA probe BF-35.

SQ Sequence 1673 BP; 528 A; 408 C; 249 G; 488 T;

alignment\_scores:

Quality: 41.50 Length: 18

Ratio: 2.964 Gaps: 1

Percent Similarity: 77.778 Percent Identity: 55.556

alignment\_block:

US-08-653-294-22 x X22323 ..

Align seg 1/1 to: X22323 from: 1 to: 1673

CC Spirochetes are pathogenic in humans and Borrelia causes epidemic and  
CC endemic relapsing fever, and Lyme borreliosis, more commonly known as  
CC Lyme disease.  
SQ Sequence 116277 BP: 42656 A; 19868 C; 14490 G; 39250 T;

alignment_scores:		
Quality:	41.50	Length: 16
Ratio:	2.964	Gaps: 1
Percent Similarity:	87.500	Percent Identity: 56.250

Align seq 1/1 to reverse of: X20249 from: 1 to: 116277

2 TyrLysLeuValIleLysIleAsnAsnIleArgIleValValLysPhe  
:::|||||::|| :::|||||:::|||::: ||  
65973 CATAAATGATTATT...TTAAATAATATAAAAATTGTTTACCTTT

eg name: N Geneseq 36:V76754

```
seq_documentation_block:
ID      V76754 standard; DNA; 264 BP.
```

16-MAR-1999 (first entry)  
Staphylococcus aureus contig SEQ ID #2443.  
Computer readable medium; vaccine; S.aureus infection; immunodetection;  
cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;  
skin infection; surgical wound infection; scalded skin syndrome;  
toxic shock syndrome; ds.

```

US-08-653-294-22 x X20248_04      ..
Align seg 1/1 to: X20248_04 from: 1 to: 110000

2 TyrLysLeuValIleLysIleAsnIleArgIleValValLysPhe 17
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
89154 CATAACTGATTTT...TTAAATAATATAAAAATGTTTTCACCTTT 89198

```

2 TyrLysLeuValIleLysIleAsnAsnIleArgIleValIleValLysPhe 17  
 :::::::::::::::::::::  
 89154 CATAAACTGATTATT...TTAAATAATATAAAATTTGTTTTACCTTTT 89

WPI; 97-3/4922/33.  
PT Polynucleotide(s) and proteins derived from *Staphylococcus aureus*  
PT stored on computer readable medium and used in the production of

stored on computer readable medium and used in the production of anti-S.aureus vaccines

Claim 1; Page 2240; 3271pp; English.

This sequence represents one of 5191 Staphylococcus aureus DNA sequences of the invention. The DNA sequences are recorded on a computer readable medium, preferably selected from a floppy or hard disk, random access memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using the S.aureus DNA sequences allows putative functions to be assigned so that protein-encoding or regulatory regions of commercial, therapeutic or industrial importance can be obtained. Specifically, sequences which are likely to encode antigens have been identified and these polypeptides can be used in a vaccine composition against S.aureus infection. The polypeptides can also be used in a kit for the immunodetection of S.aureus in a sample. S.aureus is implicated in numerous human diseases, including cellulitis, eyelid infections, food poisoning, osteomyelitis, skin and surgical wound infections, scalded skin syndrome, toxic shock syndrome, etc. Organisms transformed with the DNA sequences can be used for recombinant production of the polypeptides. The new DNA sequences (and their fragments) are useful as primers or probes for isolating homologues of any of the S.aureus DNA sequences contained on the computer readable medium.

Sequence 264 BP; 100 A; 34 C; 39 G; 89 T;

New isolated *Borrelia burgdorferi* nucleic acids - used to develop products for the detection, diagnosis, characterisation, prevention and therapy of infections, particularly Lyme disease  
 Claim 1; Page 672-737; 1128pp; English.  
 X20248 to X20402 represent polynucleotide sequences isolated from *Borrelia burgdorferi* (Bb). Products derived from Bb can be used for the detection, diagnosis, characterisation, prevention and therapy of Bb infections, e.g. Lyme disease. They can also be used for the production of biosynthetic products, e.g. enzymes. *Borrelia* belongs to a family of motile, spiral-shaped bacteria called Spirochetes.

cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;  
skin infection; surgical wound infection; scalded skin syndrome;  
Claim 1; Page 463-466; 2084pp; English.  
PS  
infection.



Align seg 1/1 to reverse of: V21209\_07 from: 1 to: 110000

2 TyrlsLeuValIleLysIleAsnArgIleValValLysPhe 17  
:::|||||:::|||||:::|||||:::|||||:::|||||  
94490 TTTAAATCATATATTAACACACACATCTGAATAATACTTAATTC 94443

seq\_name: N\_Geneseq\_36:V90892

seq\_documentation\_block:  
ID V90892 standard; DNA; 1198 BP.  
AC V90892:  
DT 18-FEB-1999 (first entry)  
DE Nucleotide sequence of cluster 45.  
KW Antigen; immunogenic cluster family; vaccine; gastritis; diagnosis;  
KW peptic ulcer; gastric adenocarcinoma; gastric lymphoma; ss.  
OS Helicobacter pylori.  
PN WO9849314-A2.  
PD 05-NOV-1998.  
PF 27-APR-1998; U08487.  
PR 14-OCT-1997; US-061958.  
PR 25-APR-1997; US-045107.  
PA (GENE-) GENELABS TECHNOLOGIES INC.  
PI Chow TP, Fry KE, Lim MY, McAttee CP;  
DR WPI; 99-009433/01.  
PT New Helicobacter pylori antigens and related nucleic acid sequences  
PT - useful in serological diagnosis and protective vaccines, providing  
PT long-lasting immune response  
PS Claim 24: Page 304-305: 402pp: English.  
CC The present sequence encodes a Helicobacter pylori antigenic protein  
CC that is characterised by immunoreactivity with H. pylori-positive  
CC antisera. The proteins are highly immunogenic and induce a long-lasting  
CC immune response that persists even after antimicrobial treatment. In  
CC antibody-detection assays, on sera, plasma, urine, saliva etc., they are  
CC highly sensitive and specific. The specification also describes 69  
CC previously unrecognised immunogenic cluster families. H. pylori antigens  
CC are used to detect H. pylori-specific antibodies, for diagnosing  
CC infection or to confirm eradication of infection, and in vaccines to  
CC protect against H. pylori infection and related diseases (gastritis,  
CC peptic ulcer, gastric adenocarcinoma/lymphoma).  
SQ Sequence 1198 BP; 395 A; 235 C; 221 G; 346 T;

alignment\_scores:  
Quality: 40.00 Length: 15  
Ratio: 2.857 Gaps: 0  
Percent Similarity: 93.333 Percent Identity: 33.333

alignment\_block:  
US-08-653-294-22 x V90892 ..

Align seg 1/1 to: V90892 from: 1 to: 1198

3 LysLeuValIleLysIleAsnArgIleValValLysPhe 17  
||| :::::::::::|||||:::|||||:::|||||  
394 AAATCACTTTTAAGACTAAATAGTTAGTAATATATCTTGATTT 438

seq\_documentation\_block:  
LOCUS AU033795 329 bp mRNA EST  
DEFINITION AU033795 Dictyostellium discoideum 5L (H.Urushihara) Dictyostellium  
discoideum cDNA clone SLB447, mRNA sequence.  
ACCESSION AU033795  
28-APR-1999

— 4 —



discoideum cDNA clone SSH480, mRNA sequence:

**Contact: Win**

discoideum cDNA clone SSH480, mRNA sequence:

Contact: Hideko Urushihara  
Institute of Biological Sciences  
University of Tsukuba  
3-3-10 Ten-nodai, Tsukuba, Ibaraki 305, Japan  
Email: d402hu@sakura.cc.tsukuba.ac.jp  
PROJECT = 'Dictyostelium discoidem cDNA project in Japan'.

Email: d402hu@sakura.cc.tsukuba.ac.jp  
 PROJECT = 'p1ctyostellum discoideum cdna project in Japan'.  
 FEATURES Location/Qualifiers  
 . source 1..628

```

1. 020
source
/organism="Dictyostelium discoideum"
/strain="AX4"
/db_xref="taxon:44589"
/clones="SLG109"
/clone_lib="Dictyostelium discoideum SL (H.Urushihara)"
/dev_stage="sluc"

```

```

BASE COUNT      242 a      84 c      94 g      208 t
ORIGIN
/uev_stage- sidg

alignment_scores:
    Quality:      50.00      Length:      15
    Ratio:        3.571      Gaps:      0
    Percent Similarity: 93.333      Percent Identity: 60.000

alignment_block:
    US-08-653-294-22 x AU039651      ..
    Align seq 1/1 to: AU039651      from: 1 to: 628

```

DEFINITION  
C94298 *Dictyostellium discoideum* SS (Hs000004) *Dictyostellium discoideum* cDNA clone SSK757, mRNA sequence.  
ACCESSION  
C94298  
VERSION  
C94298.1 GI:3218913  
KEYWORDS  
EST.

ORGANISM	Dictyostellium discoideum	Dictyostellium glabratum
Eukaryota; Dictyostellida; Dictyostelium.		

REFERENCE  
1. (Pages 1 to 345)  
Yoshino, R., Morio, T. and Tanaka, Y.  
AUTHORS

JOURNAL Unpublished (1997)  
 TITLE Developmental cDNA in *Dicystosellum discoideum*  
 COMMENT On Jan 14, 1998 this sequence version replaced gi:1797150.  
 Contact: Hideko Urushihara

**Institute of Biological Sciences  
University of Tsukuba**

FEATURES  
source  
1. .645  
Location/Qualifiers  
PROJECT = Dictyostelium discoideum cDNA project in Japan.  
Email: d402nu@sakura.cc.tsukuba.ac.jp  
Sakura Internet Access: 1997-09-07

```

/strain="AX4"

```

```

/clone="SSK757"

```

```

/dev_stage="slug"

```

```

ORIGIN
alignment_scores:
    Quality: 50.00
    Ratio: 3.571
    Gaps: 0
    Length: 15

```

Percent Similarity: 93.333      Percent Identity: 60.000

alignment\_block:

```

US-08-653-294-22 x C94298
Align seg 1/1 to: C94298 from: 1 to: 645

2 TyrLysLeuValIleAsnAsnIleArgIleValVallys 16
||||| :||||| :||||| :||||| :||||| :|||||
599 TACAATAAATTAATAAATAATTAACTTTTGTGTAAGA 643

seq_name: gb_est21:C92217

seq_documentation_block:
LOCUS      C92217       584 bp     mRNA          EST             12-JUL-1999
DEFINITION C92217 Dictyostelium discoideum SS (H.Urushihara) Dictyostelium
discoideum cDNA clone SSD246, mRNA sequence.
ACCESSION C92217
VERSION    C92217.1 GI:3074093
KEYWORDS   EST.
SOURCE     Dictyostelium discoideum.
ORGANISM   Eukaryota; Dictyosteliida; Dictyostelium.
REFERENCE  1 (bases 1 to 584)
AUTHORS    Morio,T., Urushihara,H., Saito,T., Ugawa,Y., Mizuno,H., Yoshida,M.,
Yoshino,R., Mitra,B.N., Pi,M., Sato,T., Takemoto,K., Yasukawa,H.,
Williams,J., Maeda,M., Takeuchi,I., Ochiai,H. and Tanaka,Y.
The Dictyostelium developmental cDNA project: generation and
analysis of expressed sequence tags from the first-finger stage of
development
JOURNAL    DNA Res. 5 (6), 335-340 (1998)
MEDLINE    99156227
COMMENT    On May 9, 1995 this sequence version replaced gi:804216.
Contact: Hideko Urushihara
Institute of Biological Sciences
University of Tsukuba
3-3-10 Ten-nodai, Tsukuba, Ibaraki 305, Japan
Email: d402hu@sakura.cc.tsukuba.ac.jp
PROJECT = 'Dictyostelium discoideum cDNA project in Japan'
POLYA-No.

FEATURES             Location/Qualifiers
source               1..584
                    /organism="Dictyostelium discoideum"
                    /strain="AX4"
                    /db_xref="taxon:44689"
                    /clone="SSD246"
                    /clone_lib="Dictyostelium discoideum SS (H.Urushihara)"
                    /dev_stage="slug"
BASE COUNT        228 a 79 c 82 g 193 t 2 others
ORIGIN

alignment_scores:
Quality: 48.00 Length: 15
Ratio: 3.692 Gaps: 0
Percent Similarity: 86.667 Percent Identity: 60.000

alignment_block:
US-08-653-294-22 x C92217
Align seg 1/1 to: C92217 from: 1 to: 584

2 TyrLysLeuValIleAsnAsnIleArgIleValVallys 16
||||| :||||| :||||| :||||| :||||| :|||||
515 TACAATAAATTAATAAATAATTAACTTTTGTGTAAGA 559

seq_name: gb_gss12:AQ359868

seq_documentation_block:
LOCUS      AQ359868       470 bp     DNA          GSS             06-MAR-1999
DEFINITION HS_5032.B1.B04.SP6E.RPC111 Human Male BAC Library Homo sapiens
genomic clone Plate=608 Col=7 Row=D, genomic survey sequence.
ACCESSION  AQ359868
VERSION     AQ359868.1 GI:4208744
KEYWORDS    GSS.
SOURCE      human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 470)
AUTHORS    Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L.
Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
99380589
CONTACT: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel.: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones may be purchased from Research Genetics (info@resgen.com).
BAC end Web Server: http://www.htsc.washington.edu
Plate: 608 row: D column: 7
Seq primer: SP6
Class: BAC ends
High quality sequence stop: 470.
Location/Qualifiers
source              1..470
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /clone="Plate=608 Col=7 Row=D"
                    /clone_lib="RPC111 Human Male BAC Library"
                    /sex="Male"
                    /cell_type="Lymphocytes"
                    /note="Vector: pBACe3.6; RPC111 Human Male BAC Library"
BASE COUNT        132 a 89 c 81 g 167 t 1 others
ORIGIN

alignment_scores:
Quality: 47.00 Length: 17
Ratio: 3.133 Gaps: 0
Percent Similarity: 88.235 Percent Identity: 47.059

alignment_block:
US-08-653-294-22 x AQ359868/rev
Align seg 1/1 to reverse of: AQ359868 from: 1 to: 470

1 SerTyrlsLeuValIleAsnAsnIleArgIleValVallysPh 17
||||||| :||||| :||||| :||||| :||||| :|||||
353 AGTTACAAGTGCGTTGTTAAATTAATAGGATAATNTTCCTGTTAAAGTA 304

17 e 17
303 T 303

seq_name: gb_est43:AV336157

seq_documentation_block:
LOCUS      AV336157       237 bp     mRNA          EST             11-NOV-1999
DEFINITION AV336157 RIKEN full-length enriched, adult male medulla oblongata
Mus musculus cDNA clone 6330582O13 3', mRNA sequence.
ACCESSION  AV336157
VERSION     AV336157.1 GI:6376209
KEYWORDS    EST.
SOURCE      house mouse.
ORGANISM     Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 237)
AUTHORS    Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Carninci,P., Endo,T.,
Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N., Hirozane,T., Horii,F.,
Ishii,Y., Ishikawa,T., Itoh,M., Izawa,M., Kadota,K., Kagawa,I.,
Kai,C., Kawai,J., Kikuchi,N., Kojima,Y., Koya,S., Kusakabe,M.,

```

```

US-08-653-294-22 x C94298
Align seg 1/1 to: C94298 from: 1 to: 645

2 TyrLysLeuValIleAsnAsnIleArgIleValVallys 16
||||| :||||| :||||| :||||| :||||| :|||||
599 TACAATAAATTAATAAATAAATAATTAAAGTTTGTGTAAGA 643

seq_name: gb_est21:C92217

seq_documentation_block:
LOCUS      C92217       584 bp    mRNA          EST             12-JUL-1999
DEFINITION C92217 Dictyostelium discoideum SS (H.Urushihara) Dictyostelium
discoideum cDNA clone SSD246, mRNA sequence.
ACCESSION C92217
VERSION    C92217.1 GI:3074093
KEYWORDS   EST.
SOURCE     Dictyostelium discoideum.
ORGANISM   Eukaryota; Dictyosteliida; Dictyostelium.
REFERENCE  1 (bases 1 to 584)
AUTHORS    Morio,T., Urushihara,H., Saito,T., Ugawa,Y., Mizuno,H., Yoshida,M.,
Yoshino,R., Mitra,B.N., Pi,M., Sato,T., Takemoto,K., Yasukawa,H.,
Williams,J., Maeda,M., Takeuchi,I., Ochiai,H. and Tanaka,Y.
The Dictyostelium developmental cDNA project: generation and
analysis of expressed sequence tags from the first-finger stage of
development
JOURNAL    DNA Res. 5 (6), 335-340 (1998)
MEDLINE    99156227
COMMENT     On May 9, 1995 this sequence version replaced gi:804216.
Contact: Hideko Urushihara
Institute of Biological Sciences
University of Tsukuba
3-3-10 Ten-nodai, Tsukuba, Ibaraki 305, Japan
Email: d402hu@sakura.cc.tsukuba.ac.jp
PROJECT = 'Dictyostelium discoideum cDNA project in Japan'
POLYA-No.

FEATURES             Location/Qualifiers
source               1..584
                    /organism="Dictyostelium discoideum"
                    /strain="AX4"
                    /db_xref="taxon:44689"
                    /clone="SSD246"
                    /clone_lib="Dictyostelium discoideum SS (H.Urushihara)"
                    /dev_stage="slug"
BASE COUNT        228 a 79 c 82 g 193 t 2 others
ORIGIN

alignment_scores:
Quality: 48.00 Length: 15
Ratio: 3.692 Gaps: 0
Percent Similarity: 86.667 Percent Identity: 60.000

alignment_block:
US-08-653-294-22 x C92217
Align seg 1/1 to: C92217 from: 1 to: 584

2 TyrLysLeuValIleAsnAsnIleArgIleValVallys 16
||||| :||||| :||||| :||||| :||||| :|||||
515 TACAATAAATTAATAAATAAATAATTAAAGTTTGTGTAAGA 559

seq_name: gb_gss12:AQ359868

seq_documentation_block:
LOCUS      AQ359868       470 bp    DNA          GSS             06-MAR-1999
DEFINITION HS_5032.B1.B04.SP6E.RPC111 Human Male BAC Library Homo sapiens
genomic clone Plate=608 Col=7 Row=D, genomic survey sequence.
ACCESSION  AQ359868
VERSION     AQ359868.1 GI:4208744
KEYWORDS    GSS.
SOURCE      human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 470)
AUTHORS    Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L.
Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
99380589
CONTACT: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel.: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones may be purchased from Research Genetics (info@resgen.com).
BAC end Web Server: http://www.htsc.washington.edu
Plate: 608 row: D column: 7
Seq primer: SP6
Class: BAC ends
High quality sequence stop: 470.
Location/Qualifiers
source              1..470
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /clone="Plate=608 Col=7 Row=D"
                    /clone_lib="RPC111 Human Male BAC Library"
                    /sex="Male"
                    /cell_type="Lymphocytes"
                    /note="Vector: pBACe3.6; RPC111 Human Male BAC Library"
BASE COUNT        132 a 89 c 81 g 167 t 1 others
ORIGIN

alignment_scores:
Quality: 47.00 Length: 17
Ratio: 3.133 Gaps: 0
Percent Similarity: 88.235 Percent Identity: 47.059

alignment_block:
US-08-653-294-22 x AQ359868/rev
Align seg 1/1 to reverse of: AQ359868 from: 1 to: 470

1 SerTyrLysLeuValIleAsnAsnIleArgIleValVallysPh 17
||||||| :||||| :||||| :||||| :||||| :|||||
353 AGTTACAAGTGCGTTCTTAATAAATAGGATAATNTCTCTGAAGTA 304

17 e 17
303 T 303

seq_name: gb_est43:AV336157

seq_documentation_block:
LOCUS      AV336157       237 bp    mRNA          EST             11-NOV-1999
DEFINITION AV336157 RIKEN full-length enriched, adult male medulla oblongata
Mus musculus cDNA clone 6330582O13 3', mRNA sequence.
ACCESSION  AV336157
VERSION     AV336157.1 GI:6376209
KEYWORDS    EST.
SOURCE      house mouse.
ORGANISM     Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 237)
AUTHORS    Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Carninci,P., Endo,T.,
Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N., Hirozane,T., Horii,F.,
Ishii,Y., Ishikawa,T., Itoh,M., Izawa,M., Kadota,K., Kagawa,I.,
Kai,C., Kawai,J., Kikuchi,N., Kojima,Y., Koya,S., Kusakabe,M.,

```

```

alignment_block:
US_08-653-294-22 x AV336157/rev ..
Align seg 1/1 to reverse of: AV336157 from: 1 to: 237
2 TyrTysLeuValIleLysTleAsnAsnIleArgIleValValLysPhe 17
|||||::: ||| ::::::::::: :::::::::::

```

[illegible]

```
seq_name: gb_est41:AV265649
seq_documentation_block:
```

LOCUS	AV265649	260 bp	mRNA	EST
DEFINITION	AV265649	RIKEN full-length enriched, adult male t		
		musculus cDNA clone 4930509N06 3', mRNA sequence		

ACCESSION	VERSION	GI
AV265649	AV265649.1	GI:6253684

KEYWORDS	EST.
SOURCE	house mouse.
ORGANISM	Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata  
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae  
1 (bases 1 to 260)

**AUTHORS**  
Konno, H., Aizawa, K., Akahira, S., Akiyama, J., Carn  
Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N., Hir

Ishii, Y., Ishikawa, T., Itoh, M., Izawa, M., Kadota, Kai, C., Kawai, J., Kikuchi, N., Kojima, Y., Koya, S., Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Od

Owa, C., Ozawa, Y., Saito, H., Sano, M., Sato, K., Shibata, Y., Shigemoto, Y., Shiraki, T., Sogabe, Y., Suzuki, Y., Suzuki, Y., Takehachi, F., Matsuo, M., Tanaka, K., and Tanaka, M. 1999. The effects of the

Tsunoda, Y., Watahiki, A., Watanabe, S., Yamamura, T., Yokota, T., Yoshiki, A., Yoshino, M., Muramatsu, M. and Suzuki, H., 1991. *Journal of Polymer Science, Part A: Polymer Chemistry*, **29**, 1471-1481.

TITLE	JOURNAL	COMMENT
RIKEN Mouse ESTs (Konno, H., et al.)	Unpublished (1999)	On Feb 18, 1999 this sequence version replaced gi

Contact: Yoshihide Hayashizaki  
Genome Exploration Research Group, Life Science T

Genome Science Laboratory  
The Institute of Physical and Chemical Research (IPCR)  
Sciences Center

3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan  
Tel: +81-298-36-9013  
Fax: +81-298-36-9098

Email: [genome-res@rtc.riken.go.jp](mailto:genome-res@rtc.riken.go.jp),  
URL: <http://genome.rtc.riken.go.jp/>

Sasaki, N., Izawa, M., Watanuki, M., Ozawa, K., Tanaka, M., Matsuura, S., Carninci, P., Muramatsu, M., Okazaki, Y., Hayashizaki, Y.

Transcriptional sequencing: A method for DNA sequencing using reverse transcriptase. Proc. Natl. Acad. Sci. U.S.A. 95 (7), 1203-1208.

Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Murakami, Y., and Hayashizaki, Y.

Automated filtration-based high-throughput plasma system. *Genome Res.* 9 (5), 463-470 (1999)

Carninci, P. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning. Methodod  
19-44 (1999)

**FEATURES**

Location/Qualifiers

Further details.

Please visit our web site (<http://genome.tcc.utoronto.ca>) for

```
source      : 1. .260
              /organism="Mus musculus"
              /strain="C57BL/6J"
```

```

/db_xref="taxon:10090"
/clone="4930509N06",
/clone_1db="brvrv_full-10090b-entiched

```

```

/clone_lib=RIRKEN full-length enriched,
(DH10B)"
/sex="male"

```

```
/tissue_type="testis"  
/dev_stage="adult"  
/lab_host="DH108"
```

[illegible]





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2000, 04:05:41 ; Search time 133.56 Seconds  
(without alignments)  
1.773 Million cell updates/sec

Title: US-08-653-294-23  
Perfect score: 50  
Sequence: 1 REDLRTLLRY 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 188963 seqs, 23686106 residues

Total number of hits satisfying chosen parameters: 188963

Minimum DB seq length: 0  
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database : A\_Geneseq\_36.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50	100.0	10	R41212	Peptide fragment o
2	50	100.0	10	R83075	HLA-B2702 CTL modu
3	50	100.0	10	R95423	HLA-B2705.75-84. C
4	50	100.0	10	W33785	Peptide B2705.75-8
5	50	100.0	17	R71442	Human HLA-B27-(62-
6	50	100.0	17	R71443	Human (Phe74)-HLA-
7	50	100.0	25	R41221	Peptide fragment o
8	50	100.0	25	R83091	HLA-B2702 CTL modu
9	50	100.0	25	R95417	HLA-B2705.60-84. C
10	50	100.0	337	P70590	Sequence of the hu
11	50	100.0	362	P70155	Immunomodulatory p
12	44	88.0	10	W47271	Human MHC 1 alpha
13	44	88.0	24	R71435	Human MHC 1 alpha
14	43	86.0	10	W07524	T-cell modulating
15	43	86.0	13	W29421	Glucose transport
16	43	86.0	16	R50266	HLA B27 hypervaria
17	43	86.0	17	R71440	Human MHC 1 and HL
18	43	86.0	17	R71425	Human MHC 1 alpha
19	43	86.0	17	R71426	Human MHC 1 alpha
20	43	86.0	17	R71433	Human MHC 1 alpha
21	43	86.0	17	W32583	MHC peptide repeat
22	43	86.0	17	W32581	MHC peptide repeat
23	43	86.0	17	W29422	Glucose transport
24	43	86.0	17	W45885	Peptide membrane b
25	43	86.0	24	R71424	Human MHC 1 alpha
26	43	86.0	24	R71427	Human MHC 1 alpha
27	43	86.0	25	R20116	MHC Class I-derive
28	43	86.0	25	R69619	MHC-I peptide DK-(
29	43	86.0	25	R71420	Human MHC 1 alpha
30	43	86.0	36	W32582	MHC peptide having
31	43	86.0	36	W32078	Peptide DS-A85. Pe
32	40	80.0	10	R83095	HLA-B2702 CTL modu
33	40	80.0	10	R95426	HLA-B2702.75-84(T)
34	40	80.0	10	W33788	Peptide B2702.75-8

35 40 80.0 20 1 R92909 HLA-B2702 CTL modu  
36 40 80.0 20 1 R92910 HLA-B2702 CTL modu  
37 40 80.0 20 1 W33792 Peptide B2702.84-7  
38 40 80.0 20 1 W33793 Peptide B2702.84-7  
39 40 80.0 24 1 R71434 Human MHC 1 alpha  
40 39 78.0 10 1 R83094 HLA-B2702 CTL modu  
41 39 78.0 10 1 R83096 HLA-B2702 CTL modu  
42 39 78.0 10 1 R95425 HLA-B2702.75-84(D)  
43 39 78.0 10 1 W07513 T-cell modulating  
44 39 78.0 10 1 W47267 Immunomodulatory p  
45 39 78.0 10 1 W47269 Immunomodulatory p

## ALIGNMENTS

## RESULT 1

R41212 ID R41212 standard; peptide; 10 AA.  
AC R41212;  
DT 15-MAR-1994 (first entry)  
DE Peptide fragment of Class I HLA peptide.  
KW Human leukocyte antigen; HLA; peptide; transplantation; neoplasia;  
KW Parasitic disease; cytotoxic T lymphocyte; modulation.  
OS Synthetic.  
PN WO9317699-A.  
PD 16-SEP-1993.  
PF 25-FEB-1993.  
PR 02-MAR-1992; US-844716.  
PA (STRD) UNIV LELAND STANFORD JUNIOR.  
PI Clayberger CA, Krensky AM;  
DR WPI; 93-303134/38.  
PT New peptide(s) based on Class I HLA antigen domains - used for  
PT modulating cytotoxic T-lymphocyte activity towards targets  
PS Claim 11, Page 54; 61pp; English.  
CC The peptide is used to modulate cytotoxic T-lymphocyte (CTL)  
CC activity, either by inhibition or stimulation. It can be used  
CC for inhibiting CTL toxicity in transplantations, for inducing CTL  
CC activity in parasitic diseases and neoplasia and in studies on viral  
CC infection. The peptide can also be used for identifying CTLs which  
CC bind to it and removing subsets of CTLs from a T-cell composition.  
CC This peptide sequence is more commonly found within larger peptide  
CC compounds of not more than 30 amino acids in length.  
SQ Sequence 10 AA;

Query Match 100.0%; Score 50; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.00083;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1' REDLRTLLRY 10  
|||||  
DB 1' REDLRTLLRY 10

## RESULT 2

R83075 ID R83075 standard; peptide; 10 AA.  
AC R83075;  
DT 16-MAY-1996 (first entry)  
DE HLA-B2702 CTL modulating peptide (B2702.75-84).  
KW Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;  
KW Immunosuppressant; graft versus host disorder; transplantation; therapy;  
KW Class I MHC; HLA-B2702.  
OS Synthetic.  
PN WO9526979-A1.  
PD 12-OCT-1995.  
PF 05-APR-1994; U04349.  
PR 05-APR-1994; US-222851.  
PA (STRD) UNIV LELAND STANFORD JUNIOR.  
PI Clayberger C, Krensky AM, Parham P;  
DR WPI; 95-358582/46.  
PT Extension of acceptance period of transplants from MHC unmatched

PT donor hosts - using Class I B75-84 MHC antigen of the recipient  
 PS host  
 CC Example 14; Page 34; 80pp; English.  
 CC R83061-R83085, R83090-R83096 and R92907-R92913 represent fragments of  
 CC Class I major histocompatibility complex (MHC) antigens. This sequence  
 CC corresponds to residues 75-84 of the alpha-1 domain of the class  
 CC I MHC HLA-B\*2702. These sequences can be used to extend the period of  
 CC acceptance by a recipient of a transplant from an MHC unmatched donor.  
 CC The peptides are administered to a patient in conjunction with a  
 CC subtherapeutic amount of an immunosuppressant. This is administered to  
 CC the patient for a limited period of time (compared to the lifetime  
 CC administration for current treatments). The peptides particularly  
 CC modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)  
 CC of the patient.  
 SQ Sequence 10 AA;

Query Match 100.0%; Score 50; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.00083;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 REDLRTLLRY 10  
 |||||  
 Db 1 REDLRTLLRY 10

RESULT 3  
 ID R95423 standard; peptide; 10 AA.  
 AC R95423;  
 DT 12-NOV-1996 (first entry)  
 DE HLA-B\*2705.75-84.  
 KW HLA; p74; alpha-helix; human-leucocyte-associated antigen; inhibitor;  
 KW T-cell lysate; membrane protein; mammal; heat shock protein; Hsc70; APC;  
 KW B cell; calcium influx; cytotoxic T lymphocyte; CTL; differentiation;  
 KW cytolysis; antigen presenting cell.  
 OS Synthetic.  
 PN WO9513288-A1.  
 PD 18-MAY-1995.  
 PF 10-NOV-1994; U12985.  
 PR 10-NOV-1993; US-150493.  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI Clayberger C, Krensky AM;  
 DR WPI: 95-194027/25.  
 PT Consns. comprising lymphoid surface membrane proteins - which may  
 PT inhibit cytolytic activity and differentiation of CTLs.  
 PS Example; Page 11; 29pp; English.  
 CC R95413, and R95415-R95431 represent palindromes and fragments of  
 CC human-leucocyte-associated antigens. This sequence represents the  
 CC HLA-B\*2705.75-84. These sequences can be used to isolate the protein p74  
 CC from a T-cell lysate. p74 is a T-cell surface membrane protein  
 CC associated with T-cell activation in mammalian T-cells, and is also  
 CC immunologically cross reactive with the heat shock protein Hsc70. p74 is  
 CC found in a limited number of cell types, but is particularly expressed on  
 CC B and T cells. p74 can be isolated by lysis of a suitable cell with an  
 CC amphoteric detergent, and then passed through an affinity column  
 CC containing a covalently bound HLA-B\*2702 palindromic peptide.  
 CC Compositions comprising the extracellular fragment of p74 combined with  
 CC HLA-B\*2702.60-84 (see R95416), induces calcium influx, and inhibits  
 CC cytotoxic T lymphocyte (CTL) differentiation or cytolysis. Candidate  
 CC compounds can be screened for their effect on the cytolytic activity of  
 CC T-cells, by combining them with the extracellular portion of p74 and  
 CC determining the amount of binding between the candidate compound and p74.  
 CC Modulation of CTL activity can be inhibited in a cellular composition  
 CC containing T-cells and antigen presenting cells (APCs), by adding to the  
 CC mix the extracellular portion of p74, in an amount sufficient to compete  
 CC with p74 for the binding of the p74 ligand.  
 SQ Sequence 10 AA;

Query Match 100.0%; Score 50; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.00083;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 REDLRTLLRY 10  
 |||||  
 Db 1 REDLRTLLRY 10

RESULT 4  
 ID W33785 standard; peptide; 10 AA.  
 AC W33785;  
 DT 19-JUN-1998 (first entry)  
 DE Peptide B2705.75-84 tested for immunomodulating activity.  
 KW Immunomodulating dimer; immunosuppressant drug; CTL activation;  
 KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;  
 KW rejection.  
 OS Synthetic.  
 OS Homo sapiens.  
 PN WO9744351-A1.  
 PD 27-NOV-1997.  
 PF 22-MAY-1997; U08689.  
 PR 24-MAY-1996; US-653294.  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI Beulow R, Clayberger C, Krensky AM;  
 DR WPI: 98-086530/08.  
 PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B  
 PT alpha-1 domain, used for preventing rejection of transplants or  
 PT treating autoimmune diseases  
 PS Example 1; Page 19; 41pp; English.  
 CC Peptides W33784-98 and W33778-9 were assayed for their immunomodulating  
 CC activity. A peptide-type compound or variant is claimed which has  
 CC immunomodulating activity, including the N-terminal acylated and/or  
 CC C-terminal amidated or esterified forms of up to 60 amino acids, where  
 CC the peptide-type compound comprises the formula: A-B where A =  
 CC (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or V; aa77 =  
 CC D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a hydrophobic or  
 CC small amino acid; aa82 = R or L; aa83 = G or R; and aa represents amino  
 CC acid. The sequence in the brackets may optionally be absent or truncated  
 CC at any peptide type bond within the brackets. The compounds comprise  
 CC amino acid sequences related to a Class I HLA-B alpha domain (positions  
 CC 79-84). They can be used to inhibit cytotoxic T-lymphocytes (CTL) from  
 CC undesirably attacking cells in a host or in vitro. They can also be  
 CC used in combination with antigenic peptides or proteins of interest to  
 CC activate CTLs. They can also inhibit the proliferation of T cells in  
 CC response to anti-CD3. The peptide can be used for preventing rejection  
 CC of transplants or for treating autoimmune diseases, e.g. diabetes.  
 CC rheumatoid arthritis and lupus erythematosus. The products can also be  
 CC used for detection and diagnosis.  
 SQ Sequence 10 AA;

Query Match 100.0%; Score 50; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.00083;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 REDLRTLLRY 10  
 |||||  
 Db 1 REDLRTLLRY 10

RESULT 5  
 ID R71442 standard; peptide; 17 AA.  
 AC R71442;  
 DT 12-OCT-1995 (first entry)  
 DE Human HLA-B\*27-(62-85) antigen derived peptide.  
 KW Human HLA-B\*27-(62-85) antigen derived peptide; cell receptor;  
 KW interaction modulation; arthritis; neoplasias; lupus erythematosus.  
 OS Homo sapiens.  
 PN WO9505189-A.  
 PD 23-FEB-1995.  
 PF 12-AUG-1994; U09189.  
 PR 12-AUG-1993; US-103416.  
 PA (REGC ) UNIV CALIFORNIA.



PI Goldstein A, Goodenow RS, Olsson L;  
 DR WPI: 95-098577/13.  
 PT Regulating cell surface receptor response - by modulating  
 PT interaction between MHC class I antigen and the cell surface  
 PT receptor  
 PS Example 4; Page 45; 103pp; English.  
 CC R71439-R71443 are human major histocompatibility complex class I  
 CC (MHC 1) alpha 1 domain and HLA derived peptides and fusion peptides.  
 CC They were used to modulate interactions between MHC 1/HLA and cell  
 CC surface receptors. Via competitive inhibition the peptides diminish  
 CC the receptors response. This feature may be useful for the treatment  
 CC of neoplasias, lupus erythematosus and arthritis.  
 SQ Sequence 17 AA;

Query Match 100.0%; Score 50; DB 1; Length 17;  
 Best Local Similarity 100.0%; Pred. No. 0.0014; Mismatches 0; Indels 0; Gaps 0;  
 Matches 10; Conservative 0;

OY 1 REDLRTLLRY 10  
 Db 7 REDLRTLLRY 16  
 |||||

RESULT 6  
 ID R71443 standard; peptide; 17 AA.  
 AC R71443;  
 DT 12-OCT-1995 (first entry)  
 DE Human [Phe74]-HLA-B27-(62-85) antigen derived peptide.  
 KW Human [Phe74]-HLA-B27-(62-85) antigen derived peptide; cell receptor;  
 KW interaction modulation; arthritis; neoplasias; lupus erythematosus.  
 OS Homo sapiens.  
 PN WO9505189-A.  
 PD 23-FEB-1995.  
 PF 12-AUG-1994; U09189.  
 PR 12-AUG-1993; US-105416.  
 PA (REGC ) UNIV CALIFORNIA.  
 PI Goldstein A, Goodenow RS, Olsson L;  
 DR WPI: 95-098577/13.  
 PT Regulating cell surface receptor response - by modulating  
 PT interaction between MHC class I antigen and the cell surface  
 PT receptor  
 PS Example 4; Page 45; 103pp; English.  
 CC R71439-R71443 are human major histocompatibility complex class I  
 CC (MHC 1) alpha 1 domain and HLA derived peptides and fusion peptides.  
 CC They were used to modulate interactions between MHC 1/HLA and cell  
 CC surface receptors. Via competitive inhibition the peptides diminish  
 CC the receptors response. This feature may be useful for the treatment  
 CC of neoplasias, lupus erythematosus and arthritis.  
 SQ Sequence 17 AA;

Query Match 100.0%; Score 50; DB 1; Length 17;  
 Best Local Similarity 100.0%; Pred. No. 0.0014;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 REDLRTLLRY 10  
 Db 7 REDLRTLLRY 16  
 |||||

RESULT 7  
 ID R41221 standard; peptide; 25 AA.  
 AC R41221;  
 DT 15-MAR-1994 (first entry)  
 DE Peptide fragment of HLA-B2705 antigen.  
 KW Human leukocyte antigen; HLA; peptide; transplantation; neoplasia;  
 KW parasitic disease; cytotoxic T lymphocyte; modulation.  
 OS Synthetic.  
 PN WO9317699-A.  
 PD 16-SEP-1993.

PF 25-FEB-1993; U01758.  
 PR 02-MAR-1992; US-844716.  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI Clayberger CA, Krensky AM,  
 DR WPI: 93-303134/38.  
 PT New peptide(s) based on Class I HLA antigen domains - used for  
 PT modulating cytotoxic T-lymphocyte activity towards targets  
 PS Example 13; Page 39; 61pp; English.  
 CC The peptide is used to modulate cytotoxic T-lymphocyte (CTL)  
 CC activity, either by inhibition or stimulation. It can be used for  
 CC inhibiting CTL toxicity in transplantations, for inducing CTL  
 CC activity in parasitic diseases and neoplasia and in studies on viral  
 CC infection. The peptide can also be used for identifying CTLs which  
 CC bind to it and removing subsets of CTLs from a T-cell composition.  
 CC This peptide is derived from the HLA-B2705 antigen and corresponds  
 CC to the amino acid positions 60-84 of that antigen.  
 SQ Sequence 25 AA;

Query Match 100.0%; Score 50; DB 1; Length 25;  
 Best Local Similarity 100.0%; Pred. No. 0.0022;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 REDLRTLLRY 10  
 Db 16 REDLRTLLRY 25  
 |||||

RESULT 8  
 ID R83091 standard; peptide; 25 AA.  
 AC R83091;  
 DT 16-MAY-1996 (first entry)  
 DE HLA-B2702 CTL modulating peptide (B2702.60-84).  
 KW Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;  
 KW immunosuppressant; graft versus host disorder; transplantation; therapy;  
 KW class I MHC; HLA-B2702.  
 OS Synthetic.  
 PN WO9526979-A1.  
 PD 12-OCT-1995.  
 PF 05-APR-1995; U04349.  
 PR 05-APR-1994; US-222851.  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI Clayberger C, Krensky AM, Parham P;  
 DR WPI: 95-358582/46.  
 PT Extension of acceptance period of transplants from MHC unmatched  
 PT donor hosts - using Class I B75-84 MHC antigen of the recipient  
 PT host  
 PS Example 13; Page 32; 80pp; English.  
 CC R83061-R83085, R83090-R83096 and R92907-R92913 represent fragments of  
 CC class I major histocompatibility complex (MHC) antigens. This sequence  
 CC corresponds to residues 60-84 of the alpha-1 domain of the class I MHC  
 CC HLA-B2702. These sequences can be used to extend the period of  
 CC acceptance by a recipient of a transplant from an MHC unmatched donor.  
 CC The peptides are administered to a patient in conjunction with  
 CC a subtherapeutic amount of an immunosuppressant. This is administered to  
 CC the patient for a limited period of time (compared to the lifetime  
 CC administration for current treatments). The peptides particularly  
 CC modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)  
 CC of the patient.  
 SQ Sequence 25 AA;

Query Match 100.0%; Score 50; DB 1; Length 25;  
 Best Local Similarity 100.0%; Pred. No. 0.0022;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 REDLRTLLRY 10  
 Db 16 REDLRTLLRY 25  
 |||||

RESULT 9

R95417  
ID R95417 standard; peptide; 25 AA.  
AC R95417;  
DT 12-NOV-1996 (first entry)  
DE HLA-B2705.60-84.  
KW HLA; p74; alpha1-helix; human-leucocyte-associated antigen; inhibitor;  
KW T-cell lysate; membrane protein; mammal; heat shock protein; Hsc70; APC;  
KW B cell; calcium influx; cytotoxic T lymphocyte; CTL; differentiation;  
KW cytolysis; antigen presenting cell.  
OS Synthetic.  
PN WO9513288-A1.  
PD 18-MAY-1995.  
PF 10-NOV-1994; UI2985.  
PR 10-NOV-1993; US-150493.  
PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
PI Clayberger C, Krensky AW;  
DR WPI: 95-194027/25.  
PT Compns. comprising lymphoid surface membrane proteins - which may  
PT inhibit cytolytic activity and differentiation of CTLs.  
PS Example; Page 9; 29pp; English.  
CC R95413, and R95415-R95431 represent palindromes and fragments of  
CC human-leucocyte-associated antigens. This sequence represents the  
CC HLA-B2705.60-84. These sequences can be used to isolate the protein p74  
CC from a T-cell lysate. p74 is a T-cell surface membrane protein  
CC associated with T-cell activation in mammalian T-cells, and is also  
CC immunologically cross reactive with the heat shock protein Hsc70. p74 is  
CC found in a limited number of cell types, but is particularly expressed on  
CC B and T cells. p74 can be isolated by lysis of a suitable cell with an  
CC amphoteric detergent, and then passed through an affinity column  
CC containing a covalently bound HLA-B2702 palindromic peptide.  
CC Compositions comprising the extracellular fragment of p74 combined with  
CC HLA-B2702.60-84 (see R95416), induces calcium influx, and inhibits  
CC cytotoxic T lymphocyte (CTL) differentiation or cytolysis. Candidate  
CC compounds can be screened for their effect on the cytolytic activity of  
CC T-cells, by combining them with the extracellular portion of p74 and  
CC determining the amount of binding between the candidate compound and p74.  
CC Modulation of CTL activity can be inhibited in a cellular composition  
CC containing T-cells and antigen presenting cells (APCs), by adding to the  
CC mix the extracellular portion of p74, in an amount sufficient to compete  
CC with p74 for the binding of the p74 ligand.  
SQ Sequence 25 AA;

Query Match 100.0%; Score 50; DB 1; Length 25;  
Best Local Similarity 100.0%; Pred. No. 0.0022;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 REDLRTLLRY 10  
DB 16 REDLRTLLRY 25

RESULT 10  
P70590  
ID P70590 standard; protein; 337 AA.  
AC P70590;  
DT 10-APR-1991 (first entry)  
DE Sequence of the human histocompatibility antigen HLA B27.  
KW Rheumatic disorder; genetic screening; diagnosis;  
KW ankylosing spondylitis.  
OS Homo sapiens.  
PN DE3542024-A.  
PD 04-JUN-1987.  
PF 28-NOV-1985; 542024.  
PR 28-NOV-1985; DE-542024.  
PR 21-DEC-1985; DE-545576.  
PA (BEHW ) BEHRINGERWERKE AG.  
PI Riethmuller G, Meo T, Weiss E, Szots H;  
DR WPI: 87-157893/23.  
DR N-PSDB: N70935.  
PT DNA coding for antigen HLA B27 - and diagnostic reagents contg.  
PT such DNA, antigen or antibody  
PS Disclosure; Page 5; 5pp; German.

CC The DNA may be used as a hybridisation probe for detecting the HLA  
CC B27 gene, eg for assessing susceptibility to rheumatic disorders  
CC such as ankylosing spondylitis, or may be used to transform cells  
CC for prodn. of HLA B27. The HLA B27 may be used to detect HLA B27  
CC antibody in human serum, or to produce mono- or polyclonal HLA B27  
CC antibodies for use in immunoassay.  
SQ Sequence 337 AA;

Query Match 100.0%; Score 50; DB 1; Length 337;  
Best Local Similarity 100.0%; Pred. No. 0.033;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 REDLRTLLRY 10  
DB 75 REDLRTLLRY 84

RESULT 11  
P70155  
ID P70155 standard; protein; 362 AA.  
AC P70155;  
DT 10-MAR-1993 (revised)  
DT 03-APR-1991 (first entry)  
DE Sequence encoded by genomic DNA encoding human histocompatibility  
DE antigen HLA-B 27.  
KW Ankylosing spondylitis; rheumatic disorder; diagnosis.  
OS Homo sapiens.  
PN EP-226069-A.  
PD 24-JUN-1987.  
PF 21-NOV-1986; 116139.  
PR 01-JAN-1985; DE-542024.  
PR 21-DEC-1985; DE-545576.  
PA (BEHW ) BEHRINGERWERKE AG.  
PI Szots H, Weiss E, Dorner C, Lang M, Meo T, Riethmuller G;  
DR WPI: 87-171469/25.  
DR N-PSDB: N70225.  
PT DNA coding for human histocompatibility antigen HLA-B 27 - useful  
PT for diagnosis and antigen and antibody prodn.  
PS Disclosure; p6; 13pp; German.  
CC The DNA may be used to detect the HLA-B 27 gene (opt. mutated) in  
CC human genetic material. The HLA-B 27 may be used to detect anti-HLA-  
CC B 27 antibodies in human serum. The antibodies may be used to  
CC determine HLA-B 27 levels in human serum, eg for diagnosis of  
CC rheumatic disorders, esp. ankylosing spondylitis.  
SQ Sequence 362 AA;

Query Match 100.0%; Score 50; DB 1; Length 362;  
Best Local Similarity 100.0%; Pred. No. 0.035;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 REDLRTLLRY 10  
DB 99 REDLRTLLRY 108

RESULT 12  
W47271  
ID W47271 standard; peptide; 10 AA.  
AC W47271;  
DT 22-MAY-1998 (first entry)  
DE Immunomodulatory peptide.  
KW Immunomodulator; Class I HLA-B alpha-1 domain; inhibition;  
KW transplamt rejection; treatment; autoimmune disease.  
OS Homo sapiens.  
OS Synthetic.  
FH Key  
FH Key Location/Qualifiers  
FT Misc\_difference 1..10 /note= "at least one of the amino acids is the  
FT D-isomer  
PN WO9744052-A1.  
PD 27-NOV-1997.

PF 23-APR-1997; U06705.  
 PR 22-MAY-1996; US-651650.  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI Clayberger C. Krensky AM.  
 DR WPI: 98-018220/02.  
 PT Novel immunomodulatory peptide-type compound - useful for inhibiting  
 PT transplant rejection  
 PS Claim 10; Page 36; 41pp; English.  
 CC The present sequence is an immunomodulatory peptide, which  
 CC comprises a Class I HLA-B alpha-1 domain sequence. It can be used  
 CC in a pharmaceutical composition together with a subtherapeutic dose  
 CC of an immunosuppressant, to extend the period of acceptance of a  
 CC transplant from a major histocompatibility complex (MHC) unmatched  
 CC donor, i.e. to inhibit transplant rejection. It can also be used in  
 CC the treatment of autoimmune diseases.  
 CC Peptides using the D-form amino acids are more effective  
 CC immunomodulators than their diastereomers or enantiomers.  
 SQ Sequence 10 AA;

Query Match 88.0%; Score 44; DB 1; Length 10;  
 Best Local Similarity 90.0%; Pred. No. 0.012; Mismatches 1; Indels 0; Gaps 0;  
 Matches 9; Conservative 0;

QY 1 REDLRTLLRY 10  
 | | | | |  
 Db 1 REDLRTLLRY 10

RESULT 13  
 R71435  
 ID R71435 standard; peptide: 24 AA.  
 AC R71435;  
 DT 12-OCT-1995 (first entry)  
 DE Human MHC 1 alpha 1 domain peptide [Ala69, 76]-Dk-(62-85).  
 KW Major histocompatibility complex class I; MHC 1; cell receptor;  
 KW alpha 1 domain; peptide [Ala69, 76]-Dk-(62-85); interaction modulation;  
 KW arthritis; neoplasias; lupus erythematosus.  
 OS Homo sapiens.  
 PN W09505189-A.  
 PD 23-FEB-1995.  
 PF 12-AUG-1994; U09189.  
 PR 12-AUG-1993; US-105416.  
 PA (REGC ) UNIV CALIFORNIA.  
 PI Goldstein A. Goodenow RS, Olsson L;  
 DR WPI: 95-098577/13  
 PT Regulating cell surface receptor response - by modulating  
 PT interaction between MHC class I antigen and the cell surface  
 PT receptor  
 PS Example 4; Page 45; 103pp; English.  
 CC R71424-R71438 are human major histocompatibility complex class 1  
 CC (MHC 1) alpha 1 domain derived peptides and peptide mutants, they  
 CC were used to modulate interactions between MHC 1 and cell surface  
 CC receptors. Via competitive inhibition the peptides diminish the  
 CC receptors response, this feature may be useful for the treatment  
 CC of neoplasias, lupus erythematosus and arthritis.  
 SQ Sequence 24 AA;

Query Match 88.0%; Score 44; DB 1; Length 24;  
 Best Local Similarity 90.0%; Pred. No. 0.029; Mismatches 1; Indels 0; Gaps 0;  
 Matches 9; Conservative 0;

QY 1 REDLRTLLRY 10  
 | | | | |  
 Db 14 RADLRTLLRY 23

RESULT 14  
 W07524  
 ID W07524 standard; peptide: 10 AA.  
 AC... W07524;  
 DT 04-AUG-1997 (first entry)

DE T-cell modulating peptide Dk.  
 KW T-cell modulator; autoimmune disease; tissue destruction; alpha-1 domain;  
 KW mammal; major histocompatibility complex; MHC class I; antigen; perforin;  
 KW insulin-dependent diabetes mellitus; multiple sclerosis; inflammation;  
 KW rheumatoid arthritis; psoriasis; pemphigus vulgaris; Sjogren's disease;  
 KW thyroid disease; Hashimoto's thyroiditis; myasthenia gravis; granzyme;  
 KW autologous target cell; cytokine release; T cell activation; therapy.  
 OS Synthetic.  
 PN W09635443-A1.  
 PD 14-NOV-1996.  
 PF 05-APR-1996; U04710.  
 PR 12-MAY-1995; US-440504.  
 PA (SANG-) SANGSTAT MEDICAL CORP.  
 PI Buelow R;  
 DR WPI: 96-518410/51.  
 PT Treatment of auto-immune disease by admin. of peptide(s) corresp. to  
 PT major histocompatibility complex antigens - esp. for delaying onset  
 PT of clinical symptoms of insulin dependent diabetes by modulating T  
 PT cell mediated attack on target cells  
 PS Example 1; Page 12; 24pp; English.  
 CC W07521-W07524, and W07527 represent T-cell modulating peptides that can  
 CC be used in the method of the invention. These sequences are based on a  
 CC portion of the generic peptide corresponding to residues 70-91 of the  
 CC alpha-1 domain of the major histocompatibility complex (MHC) class I  
 CC antigen (see W07510). The method is for affecting the course of an  
 CC autoimmune disease involving T-cell mediated destruction of tissue in  
 CC mammals. These peptides are used especially to treat insulin-dependent  
 CC diabetes mellitus, preferably being administered during the pre-clinical  
 CC stage to delay onset of the disease. Other diseases that can be treated  
 CC are multiple sclerosis, rheumatoid arthritis, psoriasis, pemphigus  
 CC vulgaris, Sjogren's disease, thyroid disease, Hashimoto's thyroiditis,  
 CC myasthenia gravis, etc. The peptides modulate T-cell mediated attack on  
 CC autologous target cells, and may also reduce inflammation, swelling, and  
 CC release of cytokines, perforins, granzymes etc. associated with T cell  
 CC activation.  
 SQ Sequence 10 AA;

Query Match 86.0%; Score 43; DB 1; Length 10;  
 Best Local Similarity 90.0%; Pred. No. 0.018; Mismatches 1; Indels 0; Gaps 0;  
 Matches 9; Conservative 0;

QY 1 REDLRTLLRY 10  
 | | | | |  
 Db 1 RVDLRTLLRY 10

RESULT 15  
 W29421

ID W29421 standard; peptide: 13 AA.  
 AC W29421;  
 DT 25-FEB-1998 (first entry)  
 DE Glucose transport enhancing peptide, generic formula.  
 KW Glucose transport; major histocompatibility complex class I;  
 KW MHC; antigen; diabetes mellitus; diagnostic reagent; insulin receptor;  
 KW adrenergic receptor; IGF-I receptor; transferrin.  
 OS Synthetic.  
 PN Key  
 FT Key Location/Qualifiers  
 FT Modified\_site 1  
 FT /note- Ser is optionally modified by a peptide chain  
 FT of 1-12 natural aminoacid residues "

Query Match 86.0%; Score 43; DB 1; Length 10;  
 Best Local Similarity 90.0%; Pred. No. 0.018; Mismatches 1; Indels 0; Gaps 0;  
 Matches 9; Conservative 0;

QY 1 REDLRTLLRY 10  
 | | | | |  
 Db 1 RVDLRTLLRY 10

RESULT 14  
 W07524  
 ID W07524 standard; peptide: 10 AA.  
 AC... W07524;  
 DT 04-AUG-1997 (first entry)

1-3 CN, 1-3 SH, 1-3 sulphinyl, 1-3 sulphonyl,  
 1-3 sulphoxyl, 1-3 OH, 1-3 COOH, 1-3 haloalkyl,  
 1-3 NR6R7, 1-3 alkylene-NR6R7 and 1 or 2 oxo;  
 RI=H; R2=OH.

US5516642-A.  
 14-MAY-1996.  
 16-NOV-1992: 976872.  
 16-NOV-1992: US-976872.  
 (BRIM ) BRISTOL-MYERS SQUIBB CO.  
 PI Mapelli C, Meyers CA;  
 DR WPI; 96-251003/25.  
 PT New glucose transport enhancing poly:peptide compounds - derived from  
 PT major histocompatibility complex Class I antigen, used for treating  
 PT diabetes mellitus or as diagnostic reagent.  
 PS Claim 1: Columns 11-12; 6pp; English.  
 CC The present sequence represents the generic formula of new glucose  
 CC transport enhancing peptides, derived from the major histocompatibility  
 CC complex (MHC) Class I antigen. They can be used for treating diabetes  
 CC mellitus and in screening for compounds with a similar mode of  
 CC action. They enhance or reduce the physiological response of a  
 CC cell as a result of binding to certain cellular components, such as  
 CC receptors, transporters, receptor-bound ligands or other membrane-bound  
 CC components. Typically, the peptides (Optionally covalently bound to  
 CC insulin or its functional segment) enhance glucose uptake in a  
 CC cell. They may also be used diagnostically as a ligand to determine the  
 CC presence of the cellular components described above. These peptides have  
 CC increased potency and stability in bioassay buffers compared with known  
 CC MHC Class I-derived peptides. They can be radioiodinated with gel  
 CC retention of activity, and are not subject to aggregation and gel  
 CC formation.  
 SQ Sequence 13 AA;

Query Match 86.08; Score 43; DB 1; Length 13;  
 Best Local Similarity 90.08; Pred. NO. 0.024;  
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 REDLRTLLRY 10  
 | | | | | | | |  
 Db 3 RVDLRTLLRY 12

Search completed: February 8, 2000, 04:05:42  
 Job time: 9359 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 7, 2000, 18:04:36 ; Search time 111.22 seconds  
(without alignments)  
4.241 Million cell updates/sec

Title: US-08-653-294-23  
Perfect score: 50  
Sequence: 1 REDLRTLRY 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 142080 seqs, 47169319 residues

Total number of hits satisfying chosen parameters: 142080

Minimum DB seq length: 0  
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database : PIR\_62:\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50	100.0	338	2	MHC HLA-B27-HS - h
2	50	100.0	362	2	HLHUB2
3	50	100.0	362	2	MHC class I histoc
4	50	100.0	362	2	MHC class I histoc
5	50	100.0	362	2	MHC class I histoc
6	50	100.0	362	2	human lymphocyte a
7	45	90.0	137	2	MHC HLA-B27d - hum
8	44	88.0	362	2	MHC class I histoc
9	43	86.0	328	2	MHC H-2K transplan
10	43	86.0	362	2	MHC H-2D-k protein
11	43	86.0	368	2	MHC H-2D-k w28 prote
12	40	80.0	355	2	HLA-B alpha-chain
13	40	80.0	362	2	class I histocompa
14	40	80.0	362	2	MHC class I histoc
15	40	80.0	362	2	MHC HLA-B44.2 chai
16	40	80.0	362	2	MHC class I histoc
17	39	78.0	362	2	MHC class I lympho
18	38	76.0	218	2	MHC class I HLA-J
19	38	76.0	298	1	HLMS1
20	37	74.0	367	2	C31651
21	37	74.0	426	2	C72515
22	37	74.0	1613	2	G64488
23	36	72.0	288	2	B45897
24	36	72.0	337	2	I70694
25	36	72.0	361	2	I48160
26	36	72.0	362	2	JH0291
27	36	72.0	362	2	JH0288
28	36	72.0	362	2	JH0292
29	36	72.0	369	1	HLMSKB
30	36	72.0	369	2	I56077

31 36 72.0 521 2 T16692  
32 36 72.0 930 2 B72537  
33 35 70.0 227 2 T10510  
34 35 70.0 284 2 A45840  
35 35 70.0 359 1 HLHUB4  
36 35 70.0 692 2 E70571  
37 35 70.0 1215 2 B72029  
38 35 70.0 4644 1 A38905  
39 34 68.0 157 2 G71729  
40 34 68.0 273 2 T38509  
41 34 68.0 274 2 I54483  
42 34 68.0 315 2 A35452  
43 34 68.0 319 2 G70373  
44 34 68.0 354 2 I59308  
45 34 68.0 354 2 I80168

## ALIGNMENTS

## RESULT 1

I56116  
MHC HLA-B27-HS - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 23-Jul-1999  
C:Accession: I56116  
R:Choo, S.Y.; Fan, L.A.; Hansen, J.A.  
J. Immunol. 147, 174-180, 1991  
A:Title: A novel HLA-B27 allele maps B27 allospecificity to the region around positio  
A:Reference number: I56116; MUID:91268545  
A:Accession: I56116  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-338 <RES>  
A:Cross-references: GB:M62852; NID:g187760; PIDN:AAA59647.1; PID:g187761  
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 100.0% Score 50; DB 2; Length 338;  
Best Local Similarity 100.0%; Pred. No. 0.039;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 REDLRTLRY 10  
Db 75 REDLRTLRY 84

## RESULT 2

HLHUB2  
MHC class I histocompatibility antigen HLA-B27 alpha chain precursor - human  
C:Species: Homo sapiens (man)  
C:Date: 13-Aug-1986 #sequence\_revision 28-Apr-1995 #text\_change 22-Jun-1999  
C:Accession: S07441; A25092; B25092; A94087; S34180; S44942; A90493; B24741; I55965;  
R:Weiss, E.H.; Kuon, W.; Doerner, C.; Lang, M.; Rietmueller, G.  
Immunobiology 170, 367-380, 1985  
A:Title: Organization, sequence and expression of the HLA-B27 gene: a molecular appro  
A:Reference number: S07441; MUID:86138405  
A:Accession: S07441  
A:Molecule type: DNA  
A:Residues: 1-362 <WEI>  
A:Cross-references: EMBL:X03945  
A:Note: the authors translated the codon GAC for residue 61 as Ala and the codon CAG  
A:Note: this allele is designated B\*27052 (formerly 27W)  
R:Seemann, G.H.A.; Rein, R.S.; Brown, C.S.; Ploegh, H.L.  
EMBO J. 5, 547-552, 1986  
A:Title: Gene conversion-like mechanisms may generate polymorphism in human class I g  
A:Reference number: A91061; MUID:86220133  
A:Accession: A25092  
A:Molecule type: DNA  
A:Residues: 1-362 <SEE>  
A:Cross-references: GB:X03665; NID:g32250; PIDN:CAA27302.1; PID:g871297  
A:Note: this allele is designated B\*27051 (formerly 27W)  
A:Accession: B25092

hypothetical prote  
hypothetical prote  
cby protein - Rho  
MHC class I histoc  
MHC class I histoc  
probable thrs prot  
swi/snf family hel  
dynein heavy chain  
hypothetical prote  
MHC class I histoc  
MHC HLA-B38 chain  
aldehyde reductase  
Mg(2+) and Co(2+)  
class I histocompa  
class I histocompa

A:Molecule type: DNA  
A:Residues: 1-100,'N','102-103','IA','106-362 <SE>  
A:Cross-references: GB:X03664; NID:G32236; PIDN:CRA27301.1; PID:g871296  
A:Note: this allele is designated B\*2702 (formerly 27K)  
R:Szoets, H.; Riettmueller, G.; Weiss, E.; Meo, T.  
Proc. Natl. Acad. Sci. U.S.A. 83, 1428-1432, 1986  
A:Title: Complete sequence of HLA-B\*27 cDNA identified through the characterization of s  
A:Reference number: A94087; MUID:86149317  
A:Accession: A94087  
A:Molecule type: mRNA  
A:Residues: 25-205,'V','207-362 <SZO>  
A:Cross-references: GB:M12678  
A:Note: this allele is designated B\*27052 (formerly 27W)  
R:Vilches, C.  
submitted to the EMBL Data Library, June 1993  
A:Reference number: S34180  
A:Accession: S34180  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-100,'S','102-137','D','139','Y','141-175','E','177-362 <VIL>  
A:Cross-references: EMBL:X73578  
A:Note: this allele is designated B\*2706  
R:D'Amato, M.; Sorrentino, R.  
submitted to the EMBL Data Library, May 1994  
A:Description: Identification of a novel HLA-B\*27 subtype by restriction analysis of a c  
A:Reference number: S44942  
A:Accession: S44942  
A:Molecule type: mRNA  
A:Residues: 1-139,'H','141-362 <DAM>  
A:Cross-references: EMBL:Z33453; NID:g486652; PIDN:CAA83876.1; PID:g486653  
R:Esquerria, A.; Bragado, R.; Vega, M.A.; Strominger, J.L.; Woody, J.; Lopez de Castro, J  
Biochemistry 24, 1733-1741, 1985  
A:Title: Primary structure of papain-solubilized human histocompatibility antigen HLA-B2  
A:Reference number: A90493; MUID:85226361  
A:Accession: A90493  
A:Molecule type: protein  
A:Residues: 25-265,'E','267-295 <EZQ>  
R:Vega, M.A.; Esquerria, A.; Rojo, S.; Aparicio, P.; Bragado, R.; Lopez de Castro, J.A.  
Proc. Natl. Acad. Sci. U.S.A. 82, 7394-7398, 1985  
A:Title: Structural analysis of an HLA-B\*27 functional variant: identification of residue  
A:Reference number: A94070; MUID:86042671  
A:Accession: B24741  
A:Molecule type: protein  
A:Residues: 86-100,'N','102-103','IA','106-107:171-181 <VEG>  
R:Coppin, H.L.; McDevitt, H.O.  
J. Immunol. 137, 2168-2172, 1986  
A:Title: Absence of polymorphism between HLA-B\*27 genomic exon sequences isolated from no  
A:Reference number: I55965; MUID:87009855  
A:Accession: I55965  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 25-298 <RES>  
A:Cross-references: GB:M14013; NID:gl87743; PIDN:AAAS9643.1; PID:gl87744  
R:Blasczyk, R.; Weber, M.; Salama, A.  
submitted to the EMBL Data Library, January 1995  
A:Reference number: S52291  
A:Accession: S52291  
A:Molecule type: DNA  
A:Residues: 116-192 <BLA>  
A:Cross-references: EMBL:X83737  
C:Comment: This allele for HLA-B correlates with the development of ankylosing spondylit  
C:Genetics:  
A:Gene: GDB:HLA-B  
A:Cross-references: GDB:120048; OMIM:142830  
A:Map position: 6p21.3-6p21.3  
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology  
F:1-24/Domain: intracellular #status predicted <SIG>  
F:1-24/Domain: signal sequence #status predicted <SIG>  
F:25-362/Product: class I histocompatibility antigen HLA-B\*27 alpha chain #status predict  
F:25-307/Domain: extracellular #status predicted <EXI>  
F:25-114/Domain: alpha-1 <EXI>  
F:115-206/Domain: alpha-2 <EX2>

F:220-285/Domain: immunoglobulin homology <IMM>  
F:308-331/Domain: transmembrane #status predicted <TM>  
F:332-362/Domain: intracellular #status predicted <INT>  
F:110/Binding site: carbohydrate (Asn) (covalent) #status experimental  
F:125-188/227-283/disulfide bonds: #status experimental  
Query Match 100.0%; Score 50; DB 1; Length 362;  
Best Local Similarity 100.0%; Pred. No. 0.042; 0; Indels 0; Gaps 0;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 REDLRTLLRY 10  
|||||  
Db 99 REDLRTLLRY 108  
RESULT 3  
MHC class I histocompatibility antigen HLA-B\*37 alpha chain precursor - human  
C:Species: Homo sapiens (man)  
C:Date: 16-Nov-1990 #sequence\_revision 13-Jan-1993 #text\_change 23-Jul-1999  
C:Accession: C35997  
R:Ennis, P.D.; Zemmour, J.; Salter, R.D.; Parham, P.  
Proc. Natl. Acad. Sci. U.S.A. 87, 2833-2837, 1990  
A:Title: Rapid cloning of HLA-A,B cDNA by using the polymerase chain reaction: freque  
A:Reference number: A35997; MUID:90207291  
A:Accession: C35997  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-362 <ENN>  
A:Cross-references: GB:M32320; NID:gl87792; PIDN:AAA36233.1; PID:g307224  
C:Genetics:  
A:Gene: GDB:HLA-B  
A:Cross-references: GDB:120048; OMIM:142830  
A:Map position: 6p21.3-6p21.3  
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology  
C:Keywords: transmembrane protein  
F:220-285/Domain: immunoglobulin homology <IMM>  
Query Match 100.0%; Score 50; DB 2; Length 362;  
Best Local Similarity 100.0%; Pred. No. 0.042; 0; Indels 0; Gaps 0;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 REDLRTLLRY 10  
|||||  
Db 99 REDLRTLLRY 108  
RESULT 4  
MHC class I histocompatibility antigen HLA-B\*47 precursor - human  
C:Species: Homo sapiens (man)  
C:Date: 07-Jun-1996 #sequence\_revision 07-Jun-1996 #text\_change 23-Jul-1999  
C:Accession: I68724  
R:Zemmour, J.; Ennis, P.D.; Parham, P.; Dupont, B.  
Immunogenetics 27, 281-287, 1988  
A:Title: Comparison of the structure of HLA-B\*47 to HLA-B\*13 and its relationship to 2  
A:Reference number: I54442; MUID:88152906  
A:Accession: I68724  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-362 <EM>  
A:Cross-references: GB:M19756; NID:gl84171; PIDN:AAA52664.1; PID:g386776  
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology  
Query Match 100.0%; Score 50; DB 2; Length 362;  
Best Local Similarity 100.0%; Pred. No. 0.042; 0; Indels 0; Gaps 0;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 REDLRTLLRY 10  
|||||

Db 99 REDLRTLRY 108

RESULT 5

I37485  
human lymphocyte antigen HLA-B\*27 - human  
C:Species: Homo sapiens (man)  
C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 23-Jul-1999  
C:Accession: I37485  
R:Del Porto, P.; D'Amato, M.; Fiorillo, M.T.; Tuosto, L.; Piccolella, E.; Sorrentino, R.  
J. Immunol. 153: 3093-3100, 1994  
A:Title: Identification of a novel HLA-B\*27 subtype by restriction analysis of a cytotoxic T lymphocyte antigen HLA-B\*27  
A:Reference number: I37485; MUID:94375872  
A:Accession: I37485  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-362 <RES>  
A:Cross-references: EMBL:Z33453; NID:g486652; PIDN:CAA83876.1; PID:g486653  
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 100.0%; Score 50; DB 2; Length 362;  
Best Local Similarity 100.0%; Pred. No. 0.042;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 REDLRTLRY 10

Db 99 REDLRTLRY 108

RESULT 6

I34289  
MHC HLA-B\*27d - human  
C:Species: Homo sapiens (man)  
C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 23-Jul-1999  
C:Accession: I34289  
R:Choo, S.Y.; St. John, T.; Orr, H.T.; Hansen, J.A.  
Hum. Immunol. 21: 209-219, 1988  
A:Title: Molecular analysis of the variant alloantigen HLA-B\*27d (HLA-B\*2703) identifies  
A:Reference number: I34289; MUID:88227491  
A:Accession: I34289  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-362 <RES>  
A:Cross-references: GB:M54883; NID:g187663; PIDN:AAA59616.1; PID:g187664  
C:Genetics:  
A:Gene: GDB:HLA-B  
A:Cross-references: GDB:120048; OMIM:142830  
A:Map position: 6p21.3-6p21.3  
A:Introns: 25/1; 115/1; 207/1; 299/1; 338/1; 349/1  
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 100.0%; Score 50; DB 2; Length 362;  
Best Local Similarity 100.0%; Pred. No. 0.042;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 REDLRTLRY 10

Db 99 REDLRTLRY 108

RESULT 7

I80174  
class I histocompatibility antigen - chimpanzee (fragment)  
C:Species: Pan troglodytes (chimpanzee)  
C:Date: 24-May-1996 #sequence\_revision 24-May-1996 #text\_change 23-Jul-1999  
C:Accession: I80174  
R:McAdam, S.N.; Boyson, J.E.; Liu, X.; Garber, T.L.; Hughes, A.L.; Bontrop, R.E.; Watkins  
Proc. Natl. Acad. Sci. U.S.A. 91: 5893-5897, 1994  
A:Title: A uniquely high level of recombination at the HLA-B locus.  
A:Reference number: I59308; MUID:94286544  
A:Accession: I80174

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA  
A:Residues: 1-137 <RES>  
A:Cross-references: EMBL:U05585; NID:g454787; PIDN:AAA50188.1; PID:g454788  
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 90.0%; Score 45; DB 2; Length 137;  
Best Local Similarity 90.0%; Pred. No. 0.14;  
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 REDLRTLRY 10

Db 40 RENLRTLRY 49

RESULT 8

I37515  
MHC class I histocompatibility antigen HLA-B\*2706 alpha chain precursor - human  
C:Species: Homo sapiens (man)  
C:Date: 04-Oct-1996 #sequence\_revision 04-Oct-1996 #text\_change 23-Jul-1999  
C:Accession: I37515  
R:Viaches, C.; de Pablo, R.; Kreisler, M.  
Immunogenetics 39: 219, 1994  
A:Title: Nucleotide sequence of HLA-B\*2706.  
A:Reference number: I37515; MUID:94102824  
A:Accession: I37515  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-362 <RES>  
A:Cross-references: EMBL:X73578; NID:g975658; PIDN:CAA51980.1; PID:g975659  
C:Genetics:  
A:Gene: GDB:HLA-B  
A:Cross-references: GDB:120048; OMIM:142830  
A:Map position: 6p21.3-6p21.3  
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 88.0%; Score 44; DB 2; Length 362;  
Best Local Similarity 90.0%; Pred. No. 0.58;  
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 REDLRTLRY 10

Db 99 RESLRTLRY 108

RESULT 9

I54414  
MHC H-2K transplantation antigen - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 23-Jul-1999  
C:Accession: I54414  
R:Morita, T.; Delarbre, C.; Kress, M.; Kourilsky, P.; Gachelin, G.  
Immunogenetics 21: 367-383, 1985  
A:Title: An H-2K gene of the t-w32 mutant at the T/t complex is a close parent of an  
A:Reference number: I54414; MUID:85206119  
A:Accession: I54414  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-328 <RES>  
A:Cross-references: GB:ML4827; NID:g199550; PIDN:AAA39656.1; PID:g387471  
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 86.0%; Score 43; DB 2; Length 328;  
Best Local Similarity 90.0%; Pred. No. 0.81;  
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 REDLRTLRY 10

Db 56 RVDLRTLRY 65

## RESULT 10

I71998  
MHC H-2D-k protein - mouse  
C:Species: Mus musculus (house mouse)  
C>Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 23-Jul-1999  
C:Accession: I71998  
R:Watts, S.; Vogel, J.M.; Harriman, W.D.; Itoh, T.; Staus, H.J.; Goodenow, R.S.  
J. Immunol. 139, 3878-3885, 1987  
A:Title: DNA sequence analysis of the C3H H-2Kk and H-2Dk loci. Evolutionary relationships  
A:Reference number: I56002; MUID:88060499  
A:Accession: I71998  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-362 <RES>  
A:Cross-references: GB:M18524; NID:gi199415; PIDN:AAAS3201.1; PID:g387453  
C:Genetics:  
C:Introns: 25/1; 115/1; 207/1; 299/1; 338/1; 349/1; 362/1  
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 86.0%; Score 43; DB 2; Length 362;  
Best Local Similarity 90.0%; Pred. No. 0.9;  
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 REDLRTLLRY 10  
| |||||  
DB 99 RVDLRTLLRY 108

## RESULT 11

I68705  
MHC H-2K-w28 protein - mouse  
C:Species: Mus musculus (house mouse)  
C>Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 23-Jul-1999  
C:Accession: I68705  
R:Morita, T.; Delarbre, C.; Kress, M.; Kourilsky, P.; Gachelin, G.  
Immunogenetics 21, 367-383, 1995  
A:Title: An H-2K gene of the t-w32 mutant at the T/t complex is a close parent of an H-2  
A:Reference number: I54414; MUID:85206119  
A:Accession: I68705  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-368 <RES>  
A:Cross-references: GB:M14825; NID:gi199552; PIDN:AAA39657.1; PID:g387472  
C:Genetics:  
A:Introns: 22/1; 112/1; 204/1; 296/1; 335/1; 346/1; 359/1  
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 86.0%; Score 43; DB 2; Length 368;  
Best Local Similarity 90.0%; Pred. No. 0.91; 1; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 REDLRTLLRY 10  
| |||||  
DB 96 RVDLRTLLRY 105

## RESULT 12

I37516  
HLA-B alpha-chain - human (fragment)  
C:Species: Homo sapiens (man)  
C>Date: 04-Oct-1996 #sequence\_revision 04-Oct-1996 #text\_change 23-Jul-1999  
C:Accession: I37516  
R:Gauchat-Feiss, D.; Breur-Vriesendorp, B.S.; Rufer, N.; Jeannet, M.; Roosnek, E.; Tiedt  
Tissue Antigens 44, 261-264, 1994  
A:Title: Sequencing of a novel functional HLA-B44 subtype differing in two residues in t  
A:Reference number: I37516; MUID:95176328  
A:Accession: I37516  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-355 <RES>

A:Cross-references: EMBL:X75953; NID:g9791007; PIDN:CAA53566.1; PID:g9791008  
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 80.0%; Score 40; DB 2; Length 355;  
Best Local Similarity 80.0%; Pred. No. 3.3;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 REDLRTLLRY 10  
| |||||  
DB 99 RENVLTALRY 108

## RESULT 13

S25415  
Class I histocompatibility antigen HLA-B\*4403 alpha chain - human  
C:Species: Homo sapiens (man)  
C>Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 23-Jul-1999  
C:Accession: S25415  
R:Flieschhauer, K.; Kernan, N.A.; Dupont, B.; Yang, S.Y.  
Tissue Antigens 37, 133-137, 1991  
A:Title: The two major subtypes of HLA-B44 differ for a single amino acid in codon 15  
A:Reference number: S25415; MUID:91335451  
A:Accession: S25415  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-362 <FLE>  
A:Cross-references: EMBL:X64366; NID:g32178; PIDN:CAA45718.1; PID:g32179  
C:Genetics:  
A:Gene: GDB:HLA-B  
A:Cross-references: GDB:120048; OMIM:142830  
A:Map position: 6p21.3-6p21.3  
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology  
C:Keywords: transmembrane protein  
F:220-285/Domain: immunoglobulin homology <IMM>

Query Match 80.0%; Score 40; DB 2; Length 362;  
Best Local Similarity 80.0%; Pred. No. 3.3;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 REDLRTLLRY 10  
| |||||  
DB 99 RENVLTALRY 108

## RESULT 14

A45850  
MHC class I histocompatibility antigen HLA-B13.1 - human  
C:Species: Homo sapiens (man)  
C>Date: 03-Jun-1993 #sequence\_revision 03-Jun-1993 #text\_change 23-Jul-1999  
C:Accession: A45850  
R:Rato, K.; Dupont, B.; Yang, S.Y.  
Immunogenetics 29, 117-120, 1989  
A:Title: Localization of nucleotide sequence which determines mongoloid subtype of HL  
A:Reference number: A45850; MUID:89122134  
A:Accession: A45850  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-362 <RAT>  
A:Cross-references: GB:M24075; NID:gi187703; PIDN:AAA59627.1; PID:g386884  
C:Genetics:  
A:Gene: GDB:HLA-B  
A:Cross-references: GDB:120048; OMIM:142830  
A:Map position: 6p21.3-6p21.3  
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology  
F:220-285/Domain: immunoglobulin homology <IMM>

Query Match 80.0%; Score 40; DB 2; Length 362;  
Best Local Similarity 80.0%; Pred. No. 3.3;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;



QY 1 REDLRTLLRY 10  
||:|||||  
Db 99 RENLRTALRY 108

RESULT 15  
I61861  
MHC HLA-B44.2 chain - human  
C:Species: Homo sapiens (man)  
C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 23-Jul-1999  
C:Accession: I61861  
R:Parham, P.; Lawlor, D.A.; Lomen, C.E.; Ennis, P.D.  
J. Immunol. 142, 3937-3950, 1989  
A:Title: Diversity and diversification of HLA-A,B,C alleles.  
A:Reference number: I36956; MUID:89235215  
A:Accession: I61861  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-362 <RES>  
A:Cross-references: GB:M24038; NID:G187811; PIDN:AAA59663.1; PID:G386900  
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 80.0%; Score 40; DB 2; Length 362;  
Best Local Similarity 80.0%; Pred. No. 3.3;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 REDLRTLLRY 10  
||:|||||  
Db 99 RENLRTALRY 108

Search completed: February 7, 2000, 18:04:37  
Job time: 22203 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2000, 01:25:57 ; Search time 68.63 Seconds  
(without alignments)  
4.352 Million cell updates/sec

Title: US-08-653-294-23  
Perfect score: 50  
Sequence: 1 REDRLTLRY 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 82229 seqs, 29864866 residues

Total number of hits satisfying chosen parameters: 82229

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database : SwissProt\_38.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50	100.0	338	1B20_HUMAN	P30467 homo sapien
2	50	100.0	361	1B14_HUMAN	P03989 homo sapien
3	50	100.0	362	1B16_HUMAN	P19373 homo sapien
4	50	100.0	362	1B18_HUMAN	P10318 homo sapien
5	50	100.0	362	1B29_HUMAN	P18463 homo sapien
6	50	100.0	362	1B45_HUMAN	P30485 homo sapien
7	44	88.0	362	1B19_HUMAN	O08136 homo sapien
8	43	86.0	328	1HA1Q_MOUSE	P14428 mus musculus
9	43	86.0	362	1HA13_MOUSE	P14426 mus musculus
10	43	86.0	368	1HA1W_MOUSE	P03991 mus musculus
11	40	80.0	362	1B05_HUMAN	P30461 homo sapien
12	40	80.0	362	1B41_HUMAN	P30481 homo sapien
13	40	80.0	362	1B42_HUMAN	P30482 homo sapien
14	38	76.0	298	1HA1Y_MOUSE	P01895 mus musculus
15	38	76.0	365	1HA1Z_MOUSE	P01900 mus musculus
16	36	72.0	369	1HA1B_MOUSE	P01901 mus musculus
17	35	70.0	227	1CBY1_MOUSE	O33513 rhodobacter
18	35	70.0	359	1B40_HUMAN	P10320 homo sapien
19	35	70.0	692	1SYT_MYCTU	O06200 mycobacteri
20	35	70.0	4644	1DYHC_RAT	P38650 rattus norv
21	34	68.0	157	1Y18_RICPR	G92dx6 rickettsia
22	34	68.0	315	1ALDR_BOVIN	P16116 bos taurus
23	34	68.0	354	1GPR2_HUMAN	P46092 homo sapien
24	34	68.0	359	1B01_PANTR	P13750 pan troglod
25	34	68.0	362	1B01_GORGO	P30379 gorilla gor
26	34	68.0	362	1B02_GORGO	P30380 gorilla gor
27	34	68.0	362	1B03_GORGO	P30381 gorilla gor
28	34	68.0	362	1B15_HUMAN	P10317 homo sapien
29	34	68.0	362	1B47_HUMAN	P30487 homo sapien
30	34	68.0	362	1B49_HUMAN	P18464 homo sapien
31	34	68.0	362	1B52_HUMAN	P30489 homo sapien
32	34	68.0	362	1B53_HUMAN	P30490 homo sapien
33	34	68.0	362	1B54_HUMAN	P30491 homo sapien
34	34	68.0	362	1B60_HUMAN	P18465 homo sapien

```

35      34      68.0      362      1      1B61_HUMAN      P30497 homo sapien
36      34      68.0      362      1      1B62_HUMAN      P10319 homo sapien
37      34      68.0      362      1      1LAH_HUMAN      P01893 homo sapien
38      34      68.0      365      1      1A23_HUMAN      P30447 homo sapien
39      34      68.0      385      1      1A24_HUMAN      P05534 homo sapien
40      34      68.0      403      1      1FLGE_HELMU      P50611 helicobacte
41      34      68.0      912      1      1MUTS_SYNY3      P73769 synechocyst
42      34      68.0      2549      1      1FRAP_HUMAN      P42345 homo sapien
43      34      68.0      2549      1      1FRAP_RAT        P42346 rattus norv
44      33      66.0      346      1      1P12_TRYBB       P23734 trypanosoma
45      33      66.0      359      1      1YQGM_BACSU      P54490 bacillus su

```

## ALIGNMENTS

```

RESULT 1
1B20_HUMAN
ID      1B20_HUMAN      STANDARD;      PRT;      338 AA.
AC      P30467;
DT      01-APR-1993 (Rel. 25, Created)
DT      01-APR-1993 (Rel. 25, Last sequence update)
DT      01-APR-1993 (Rel. 25, Last annotation update)
DE      HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-27 B*2707 ALPHA CHAIN
      (B27-HS).
GN      HLA-B OR HLAB.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC      Eutheria; Primates; Catarrhini; Hominoidea; Homo.
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE; 91268545.
RA      CHOO Y.S., FAN L.A., HANSEN J.A.;
RT      "A novel HLA-B*27 allele maps B27 allospecificity to the region around
RT      position 70 in the alpha 1 domain.";
RL      J. Immunol. 147:174-180(1991).
CC      -|- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC      THE IMMUNE SYSTEM.
CC      -|- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC      MICROGLOBULIN).
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL; M62852; AAA59647.1; -.
CC      HSSP; P03989; 1HSA.
CC      TIM; 142830; -.
CC      PROSITE; PS00290; IG_MHC; 1.
CC      PFAM; PF00047; Ig; 1.
CC      PFAM; PF00129; MHC_I; 1.
CC      MHC I; Transmembrane; Glycoprotein.
CC      DOMAIN 1 90
CC      FT DOMAIN 91 182
CC      FT DOMAIN 183 274
CC      FT DOMAIN 275 284
CC      FT TRANSMEM 285 308
CC      FT DOMAIN 309 338
CC      FT CARBOHYD 86 86
CC      FT DISULFID 101 164
CC      FT DISULFID 203 259
CC      SQ SEQUENCE 338 AA; 37804 MW; 33FB8134 CRC32;

```

```

EXTRACELLULAR ALPHA-1.
EXTRACELLULAR ALPHA-2.
EXTRACELLULAR ALPHA-3.
CONNECTING PEPTIDE.
CYTOPLASMIC TAIL.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.

```

Query Match 100.0%; Score 50; DB 1; Length 338;  
Best Local Similarity 100.0%; Pred. No. 0.013;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 REDRLTLRY 10

```

Db 75 REDRLTLRY 84
|||||
RESULT 2
ID 1B14_HUMAN STANDARD; PRT; 361 AA.
AC P03989;
DT 23-OCT-1986 (Rel. 02, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-27 ALPHA CHAIN PRECURSOR.
GN HLA-B OR HLAB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC [1]
RN SEQUENCE FROM N.A.
RP MEDLINE; 86138405.
RA WEISS E.H., KUON W., DOERNER C., LANG M., RIETHMUELLER G.;
RT "Organization, sequence and expression of the HLA-B27 gene: a
RT molecular approach to analyze HLA and disease associations.";
RL Immunobiology 170:367-380(1985).
RN [2]
RN SEQUENCE OF 25-361 FROM N.A.
RP MEDLINE; 86149317.
RA SZOETS H., RIETHMUELLER G., WEISS E., MEO T.;
RT "Complete sequence of HLA-B27 cDNA identified through the
RT characterization of structural markers unique to the HLA-A, -B, and
RT -C allelic series.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:1428-1432(1986).
RN [3]
RN SEQUENCE OF 25-295.
RX MEDLINE; 85226361.
RA EZQUERRA A., BRAGADO R., VEGA M.A., STROMINGER J.L., WOODY J.,
RA LOPEZ DE CASTRO J.A.;
RT "Primary structure of papain-solubilized human histocompatibility
RT antigen HLA-B27.";
RL Biochemistry 24:1733-1741(1985).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 25-300.
RX MEDLINE; 92405152.
RA MADDEN D.R., GORGA J.C., STROMINGER J.L., WILEY D.C.;
RT "The three-dimensional structure of HLA-B27 at 2.1-A resolution
RT suggests a general mechanism for tight peptide binding to MHC.";
RL Cell 70:1035-1048(1992).
RN [5]
RP X-RAY CRYSTALLOGRAPHY.
RX MEDLINE; 92018187.
RA MADDEN D.R., GORGA J.C., STROMINGER J.L., WILEY D.C.;
RT "The structure of HLA-B27 reveals nonamer self-peptides bound in an
RT extended conformation.";
RL Nature 353:321-325(1991).
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC THE IMMUNE SYSTEM.
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC MICROGLOBULIN).
CC -!- DISEASE: THIS PROTEIN CORRELATES WITH THE DEVELOPMENT OF
CC ANKYLOSING SPONDYLITIS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X03945; CAA27578.1; ALT_TERM.
CC PIR; A25128; HLHUB2.
CC PIR; S07441; S07441.
CC PDB; 1HSA; 15-OCT-92.
CC MIM; 142830; -

```

---

```

DR PROSITE; PS00290; IG_MHC; 1.
DR PFAM; PF00047; Ig; 1.
DR PFAM; PF00129; MHC_I; 1.
KW MHC I; Transmembrane; Glycoprotein; Signal; 3D-structure.
FT SIGNAL 1 24
FT CHAIN 25 361
FT DOMAIN 25 114
FT DOMAIN 115 206
FT DOMAIN 207 298
FT DOMAIN 299 308
FT TRANSMEM 309 332
FT DOMAIN 333 361
FT CARBOHYD 110 110
FT DISULFID 125 188
FT DISULFID 227 283
FT CONFLICT 206 206
FT CONFLICT 266 266
FT STRAND 27 38
FT TURN 39 41
FT STRAND 42 52
FT TURN 53 54
FT STRAND 55 61
FT TURN 62 63
FT STRAND 70 71
FT STRAND 74 76
FT TURN 77 78
FT HELIX 81 108
FT TURN 109 110
FT TURN 113 114
FT STRAND 118 127
FT TURN 129 130
FT STRAND 133 142
FT TURN 143 144
FT STRAND 145 150
FT TURN 152 153
FT STRAND 157 159
FT HELIX 162 173
FT TURN 174 175
FT HELIX 176 185
FT TURN 186 186
FT TURN 187 198
FT TURN 199 199
FT HELIX 200 203
FT TURN 204 204
FT STRAND 207 207
FT STRAND 210 217
FT STRAND 222 233
FT STRAND 238 243
FT TURN 244 245
FT STRAND 246 247
FT HELIX 249 251
FT STRAND 253 254
FT STRAND 258 259
FT STRAND 265 274
FT TURN 275 276
FT HELIX 278 280
FT STRAND 281 286
FT TURN 288 289
FT STRAND 294 296
FT STRAND 361 AA; 40464 MW; 802130DS CRC32;
SQ SEQUENCE

```

---

```

Query Match 100.08; Score 50; DB 1; Length 361;
Best Local Similarity 100.08; Pred. No. 0.014;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 REDRLTLRY 10
Db 99 REDRLTLRY 108
RESULT 3

```

```

1B16_HUMAN
ID 1B16_HUMAN STANDARD: PRT: 362 AA.
AC P19373;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-27 B*2703 ALPHA CHAIN
DE PRECURSOR (B-27D).
GN HLA-B OR HLAB
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 88227491.
RA CHOO S.Y., ST JOHN T., ORR H.T., HANSEN J.A.;
RT "Molecular analysis of the variant allotype HLA-B*2703)
RL Hum. Immunol. 21:209-219(1988).
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC THE IMMUNE SYSTEM.
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC MICROGLOBULIN).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M54883; AAA59616.1; .
DR HSSP; P03989; 1HSA.
DR MIM; 142830; .
DR PROSITE; PS00290; IG_MHC; 1.
DR PFAM; PF00047; Ig; 1.
DR PFAM; PF00129; MHC_I; 1.
DR MHC I; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 24
FT CHAIN 25 362 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
FT DOMAIN 25 114 B-27 B*2703 ALPHA CHAIN.
FT DOMAIN 115 206 EXTRACELLULAR ALPHA-1.
FT DOMAIN 207 298 EXTRACELLULAR ALPHA-2.
FT DOMAIN 299 308 EXTRACELLULAR ALPHA-3.
FT TRANSMEM 309 332 CONNECTING PEPTIDE.
FT DOMAIN 333 362 CYTOPLASMIC TAIL.
FT CARBOHYD 110 110 BY SIMILARITY.
FT DISULFID 125 188 BY SIMILARITY.
FT DISULFID 227 283 BY SIMILARITY.
SQ SEQUENCE 362 AA; 40402 MW; 7261C3AB CRC32;

Query Match 100.0%; Score 50; DB 1; Length 362;
Best Local Similarity 100.0%; Pred. No. 0.015;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 REDRLTLRY 10
Db 99 REDRLTLRY 108
|||||
|||||

RESULT 4
1B18_HUMAN STANDARD: PRT: 362 AA.
ID 1B18_HUMAN
AC P10318;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-27 B*2705 ALPHA CHAIN
DE... PRECURSOR (B-27W) (B27.1).
GN HLA-B OR HLAB.

```

```

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 86220133.
RA SEEMANN G.H.A., REIN R.S., BROWN C.S., PLOEGH H.L.;
RT "Gene conversion-like mechanisms may generate polymorphism in human
RT class I genes."
RL EMBO J. 5:547-552(1986).
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE; 86138405.
RA WEISS E.H., KUON W., DOERNER C., LANG M., RIETHMUELLER G.;
RT "Organization, sequence and expression of the HLA-B*27 gene: a
RT molecular approach to analyze HLA and disease associations."
RL Immunobiology 170:367-380(1985).
RN [3]
3D-STRUCTURE MODELING OF 115-206.
RX MEDLINE; 95148615.
RA ROGNAN D., SCAPOZZA L., FOLKERS G., DASER A.;
RT "Rational design of nonnatural peptides as high-affinity ligands for
RT the HLA-B*2705 human leukocyte antigen."
RL Proc. Natl. Acad. Sci. U.S.A. 92:753-757(1995).
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC THE IMMUNE SYSTEM.
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC MICROGLOBULIN).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X03665; CAA27302.1; .
DR EMBL; X03666; CAA27302.1; JOINED.
DR EMBL; M12967; AAA36221.1; .
DR PIR; A25092; HLHUBW.
DR PDB; IROG; 30-SEP-94.
DR PDB; IROH; 30-SEP-94.
DR PDB; IROI; 30-SEP-94.
DR PDB; IROJ; 30-SEP-94.
DR PDB; IROK; 30-SEP-94.
DR PDB; IROL; 30-SEP-94.
DR MIM; 142830; .
DR PROSITE; PS00290; IG_MHC; 1.
DR PFAM; PF00047; Ig; 1.
DR PFAM; PF00129; MHC_I; 1.
DR MHC I; Transmembrane; Glycoprotein; Signal; 3D-structure.
FT SIGNAL 1 24
FT CHAIN 25 362 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
FT DOMAIN 25 114 B-27 B*2705 ALPHA CHAIN.
FT DOMAIN 115 206 EXTRACELLULAR ALPHA-1.
FT DOMAIN 207 298 EXTRACELLULAR ALPHA-2.
FT DOMAIN 299 308 EXTRACELLULAR ALPHA-3.
FT TRANSMEM 309 332 CONNECTING PEPTIDE.
FT DOMAIN 333 362 CYTOPLASMIC TAIL.
FT CARBOHYD 110 110 BY SIMILARITY.
FT DISULFID 125 188 BY SIMILARITY.
FT DISULFID 227 283 BY SIMILARITY.
SQ SEQUENCE 362 AA; 40428 MW; 73243566 CRC32;

Query Match 100.0%; Score 50; DB 1; Length 362;
Best Local Similarity 100.0%; Pred. No. 0.015;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 REDRLTLRY 10
|||||

```



```

RA MORITA T., DELARBRE C., KRESS M., KOURILSKY P., GACHELIN G.;
RT "An H-2k gene of the tw32 mutant at the T/t complex is a close parent
RT of an H-2Kq gene.";
RL Immunogenetics 21:367-383(1985).
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC THE IMMUNE SYSTEM.
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC MICROGLOBULIN).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M14827; AAA39656.1; -.
DR HSSP: P01901; ZCRB.
DR MGD: MGI:95904; H2-K.
DR PROSITE: PS00290; IG_MHC; 1.
DR PFAM: PF00047; Ig; 1.
DR PFAM: PF00129; MHC_I; 1.
DR MHC I; Transmembrane; Glycoprotein.
KW NON_TER 1
FT DOMAIN <1 71 EXTRACELLULAR ALPHA-1.
FT DOMAIN 72 163 EXTRACELLULAR ALPHA-2.
FT DOMAIN 164 255 EXTRACELLULAR ALPHA-3.
FT DOMAIN 256 285 CONNECTING PEPTIDE.
FT TRANSMEM 266 289
FT DOMAIN 290 328
FT DISULFID 82 145
FT DISULFID 184 240
FT CARBOHYD 67 67
FT CARBOHYD 157 157
FT SEQUENCE 328 AA; 36855 MW; 3845B867 CRC32;
SQ
Query Match 86.0%; Score 43; DB 1; Length 328;
Best Local Similarity 90.0%; Pred. No. 0.31;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 REDLRTLLRY 10
I | | | | | | |
DB 56 RVDLRTLLRY 65
RESULT 9
HAL3_MOUSE
ID ID HAL3_MOUSE STANDARD; PRT; 362 AA.
AC AC P14426;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE H-2 CLASS I HISTOCOMPATIBILITY ANTIGEN, D-K ALPHA CHAIN PRECURSOR
DE (H-2D(K)).
GN H2-D.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
[1]
RC STRAIN=C3H;
RX MEDLINE: 88060499.
RA WATTS S., VOGEL J.M., HARRIMAN W.D., ITOH T., STAUSS H.J.,
RA GOODENOW R.S.
RT "DNA sequence analysis of the C3H H-2Kk and H-2dk loci. Evolutionary
RT relationships to H-2 genes from four other mouse strains.";
RL J. Immunol. 139:3878-3885(1987).
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC THE IMMUNE SYSTEM.
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC MICROGLOBULIN).

```

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

[illegible]

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; M14825; AAA39657.1;	-
DR EMBL; X00172; CRA24997.1;	-
DR PIR; A02197; A02197.	
DR PIR; A02196; A02196.	
DR HSP; P01901; 2CKB.	
DR MGD; MG1-95904; H2-K.	
DR PROSITE; FS00290; IG_MHC; 1.	
DR PFAM; PF00047; Ig; 1.	
DR PFAM; PF00129; MHC-I; 1.	
KW MHC I; Transmembrane; Glycoprotein; Signal.	
FT SIGNAL	1
FT FT	21
FT CHAIN	22
FT FT	368
FT FT	
FT DOMAIN	22
FT DOMAIN	111
FT DOMAIN	112
FT DOMAIN	204
FT DOMAIN	295
FT DOMAIN	305
FT TRANSMEM	306
FT TRANSMEM	329
FT DOMAIN	330
FT DISULFID	330
FT DISULFID	122
FT DISULFID	185
FT DISULFID	224
FT CARBOHYD	280
FT CARBOHYD	107
FT CARBOHYD	107
FT CARBOHYD	197
FT CARBOHYD	197
SEQUENCE	368 AA; 41103 MW; E03C7D0F CRC32;

```

Query Match      86.0%; Score 43; DB 1; Length 362;
Best Local Similarity 90.0%;
Pred. NO. 0.34;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 RDRLTLLRY 10
          | | | | | | |
Db      99 RVDRLTLLRY 108

```

```

Query Match      86.0%; Score 43; DB 1; Length 368;
Best Local Similarity 90.0%; Pred. No. 0.35;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 REDRLTILRY 10
        | | | | | | | |

```



RA LIN L., TOKUNAGA K., NAKAJIMA F., ISHIKAWA Y., KASHIWASE K.,  
 RA TANAKA H., KUMATA S., SIELTSEVA E., AKAZA T., TADOKORO K.,  
 RA SHIBATA Y., CHANDANAYONG D., JUJI T.,  
 RT "Both HLA-B\*1301 and B\*1302 exist in Asian populations and are  
 associated with different haplotypes."  
 RL Hum. Immunol. 43:51-56(1995).  
 CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO  
 THE IMMUNE SYSTEM.  
 CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
 MICROGLOBULIN).  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; M24041; AAA59660.1; -  
 DR EMBL; M19757; AAA52657.1; -  
 DR EMBL; D50291; BAA08822.1; -  
 DR HSSP; P30491; IALM.  
 DR MIM; 142830; -  
 DR PROSITE; PS00290; IG\_MHC; 1.  
 DR PFAM; PF00047; Ig; 1.  
 DR PFAM; PF00129; MHC\_I; 1.  
 KW MHC I; Transmembrane; Glycoprotein; Signal.  
 FT SIGNAL 1 24  
 FT CHAIN 25 362 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,  
 B-13 B\*1301 ALPHA CHAIN.  
 FT DOMAIN 25 114 EXTRACELLULAR ALPHA-1.  
 FT DOMAIN 115 206 EXTRACELLULAR ALPHA-2.  
 FT DOMAIN 207 298 EXTRACELLULAR ALPHA-3.  
 FT DOMAIN 299 309 CONNECTING PEPTIDE.  
 FT TRANSMEM 310 333  
 FT DOMAIN 334 362 CYTOPLASMIC TAIL.  
 FT CARBOHYD 110 110 BY SIMILARITY.  
 FT DISULFID 125 188 BY SIMILARITY.  
 FT DISULFID 227 283 BY SIMILARITY.  
 SQ SEQUENCE 362 AA; 40474 MW; 28B67875 CRC32;  
 Query Match 80.0%; Score 40; DB 1; Length 362;  
 Best Local Similarity 80.0%; Pred. No. 1.3;  
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 REDLRTLLRY 10  
 Db 99 RENLRTALRY 108  
 RESULT 12  
 ID 1B41\_HUMAN STANDARD; PRT; 362 AA.  
 AC P30481;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, BW-44(B-12) B\*4402 ALPHA CHAIN  
 DE PRECURSOR (B44.2).  
 GN HLA-B OR HLAB.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 89235215.  
 RA PARHAM P., LAWLER D.A., LOMEN C.E., ENNIS P.D.;  
 RT "Diversity and diversification of HLA-A,B,C alleles."  
 RL J. Immunol. 142:3937-3950(1989).  
 CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO  
 THE IMMUNE SYSTEM.

CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
 MICROGLOBULIN).  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; M24038; AAA59663.1; -  
 DR HSSP; P30491; IALM.  
 DR MIM; 142830; -  
 DR PROSITE; PS00290; IG\_MHC; 1.  
 DR PFAM; PF00047; Ig; 1.  
 DR PFAM; PF00129; MHC\_I; 1.  
 KW MHC I; Transmembrane; Glycoprotein; Signal.  
 FT SIGNAL 1 24  
 FT CHAIN 25 362 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,  
 BW-44(B-12) B\*4402 ALPHA CHAIN.  
 FT DOMAIN 25 114 EXTRACELLULAR ALPHA-1.  
 FT DOMAIN 115 206 EXTRACELLULAR ALPHA-2.  
 FT DOMAIN 207 298 EXTRACELLULAR ALPHA-3.  
 FT TRANSMEM 310 333  
 FT DOMAIN 334 362 CYTOPLASMIC TAIL.  
 FT CARBOHYD 110 110 BY SIMILARITY.  
 FT DISULFID 125 188 BY SIMILARITY.  
 FT DISULFID 227 283 BY SIMILARITY.  
 SQ SEQUENCE 362 AA; 40481 MW; D0AE6DD5 CRC32;  
 Query Match 80.0%; Score 40; DB 1; Length 362;  
 Best Local Similarity 80.0%; Pred. No. 1.3;  
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 REDLRTLLRY 10  
 Db 99 RENLRTALRY 108  
 RESULT 13  
 ID 1B42\_HUMAN STANDARD; PRT; 362 AA.  
 AC P30482;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, BW-44(B-12) B\*4403 ALPHA CHAIN  
 DE PRECURSOR.  
 GN HLA-B OR HLAB.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 91335451.  
 RA FLEISCHHAUER K., KERNAN N.A., DUPONT B., YANG S.Y.;  
 RT "The two major subtypes of HLA-B\*44 differ for a single amino acid in  
 codon 156."  
 RL Tissue Antigens 37:133-137(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 96435470.  
 RA ADAMS E.J., LITTLE A.M., ARNETT K.L., MCAULEY J.E., WILLIAMS R.C.,  
 RA FARHAM P.;  
 RT "Three new HLA-B alleles found in Mexican-Americans."  
 RL Tissue Antigens 46:414-416(1995).  
 CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO  
 THE IMMUNE SYSTEM.  
 CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
 MICROGLOBULIN).

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----

CC EMBL; X64366; CAA45718.1; -;  
 CC EMBL; L42282; AAB51454.1; -;  
 CC EMBL; L42283; AAB51455.1; -;  
 CC PIR; S25415; S25415.  
 CC HSP; P30491; IAIM.  
 CC MIM; 142830; -;  
 CC PROSITE; PS00290; IG\_MHC; 1.  
 CC PFAM; PF00047; Ig; 1.  
 CC PFAM; PF00129; MHC.I; 1.  
 CC MHC I; Transmembrane; Glycoprotein; Signal.  
 CC SIGNAL 1 24  
 CC CHAIN 25 362  
 CC HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,  
 CC BW-44(B-12) B\*4403 ALPHA CHAIN.  
 CC EXTRACELLULAR ALPHA-1.  
 CC DOMAIN 25 114  
 CC DOMAIN 115 206  
 CC DOMAIN 207 298  
 CC DOMAIN 299 309  
 CC DOMAIN 310 333  
 CC TRANSMEM 334 362  
 CC CYTOPLASMIC TAIL.  
 CC CARBOHYD 110 110  
 CC BY SIMILARITY.  
 CC DISULFID 125 188  
 CC BY SIMILARITY.  
 CC SEQUENCE 362 AA; 40479 MW; E72CE669 CRC32;  
 SQ

Query Match 80.0%; Score 40; DB 1; Length 362;  
 Best Local Similarity 80.0%; Pred. No. 1.3;  
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 REDLTLRLY 10  
 ||:|||||  
 Db 99 RENLRLALRY 108

RESULT 14  
 HALY\_MOUSE  
 ID HALY\_MOUSE STANDARD; PRT; 298 AA.  
 AC P01895;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 01-JAN-1990 (Rel. 13, Last annotation update)  
 DE H-2 CLASS I HISTOCOMPATIBILITY ANTIGEN, ALPHA CHAIN (CLONE PAG64)  
 DE (FRAGMENT).  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;  
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 RN [1]  
 RP SEQUENCE OF 1-285 FROM N.A. (CLONE PAG64).  
 RX MEDLINE; 84093573.  
 RA BRICKELL P.M., LATCHMAN D.S., MURPHY D., WILLISON K.,  
 RA RIGBY P.W.J.;  
 RT "Activation of a Qa/Tla class I major histocompatibility antigen gene is a general feature of oncogenesis in the mouse."  
 RL Nature 306:756-760(1983).  
 RN [2]  
 RP SEQUENCE OF 90-298 FROM N.A. (CLONE PH-2D-1).  
 RX MEDLINE; 82013602.  
 RA BREGGERE F., ABASTADO J.P., KVIST S., RASK L., LALANNE J.-L.,  
 RA GAROFF H., CAMI B., WIMAN K., LARHAMMAR D., PETERSON P.A.,  
 RA GACHELIN G., KOURILSKY P., DOBERSTEIN B.;  
 RT "Structure of C-terminal half of two H-2 antigens from cloned mRNA."  
 RL Nature 292:78-81(1981)  
 CC -!- ALTERNATIVE PRODUCTS: OF THE PRIMARY TRANSCRIPT OF THIS Qa/Tla  
 CC GENE PRODUCES MRNA WITHOUT THE NUCLEOTIDES ENCODING THE SEVENTH

CC EXON (RESIDUES 286-298).  
 CC -!- MISCELLANEOUS: THE ANTIGEN ENCODED BY THE PAG64 CLONE, NORMALLY  
 CC EXPRESSED ONLY ON LYMPHOCYTE SUBSETS AND THYMOCYTES, IS FOUND IN  
 CC INCREASED CONCENTRATIONS IN ALL TRANSFORMED CELLS TESTED, IN AN  
 CC EMBRYONIC CARCINOMA CELL LINE, AND IN PLURIPOTENT EMBRYONIC CELLS.  
 CC ACTIVATION OF THE GENE APPEARS TO BE A GENERAL FEATURE OF  
 CC CARCINOGENESIS IN THE MOUSE. THE PAG64 CLONE, AND RELATED CDNA  
 CC CLONES, HAVE IN COMMON A REPETITIVE SEQUENCE IN THEIR 3'  
 CC TRANSCRIPTION UNITS THAT HAS CHARACTERISTICS OF A TRANSPOSABLE  
 CC ELEMENT.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

CC EMBL; X00246; CAA25061.1; -;  
 CC PIR; A02194; HLMS1.  
 CC HSP; P01897; 1LD9.  
 CC PROSITE; PS00290; IG\_MHC; 1.  
 CC PFAM; PF00047; Ig; 1.  
 CC PFAM; PF00129; MHC.I; 1.  
 CC MHC I; Transmembrane; Glycoprotein; Alternative splicing.  
 CC NON\_TER 1 1  
 CC CARBOHYD 43 43  
 CC BY SIMILARITY.  
 CC CARBOHYD 133 133  
 CC POTENTIAL.  
 CC BY SIMILARITY.  
 CC DISULFID 58 121  
 CC BY SIMILARITY.  
 CC DISULFID 160 216  
 CC BY SIMILARITY.  
 CC VARIANT 212 212  
 CC Q -> L (IN REF. 2).  
 CC VARIANT 223 223  
 CC MISSING (IN REF. 2).  
 CC VARIANT 285 285  
 CC V -> G (IN REF. 2).  
 CC SEQUENCE 298 AA; 33850 MW; 56E1DA69 CRC32;  
 SQ

Query Match 76.0%; Score 38; DB 1; Length 298;  
 Best Local Similarity 80.0%; Pred. No. 2.6;  
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 REDLTLRLY 10  
 | |||||  
 Db 32 RVDLRLALRY 41

RESULT 15  
 HAL2\_MOUSE  
 ID HAL2\_MOUSE STANDARD; PRT; 365 AA.  
 AC P01900;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 13-AUG-1987 (Rel. 05, Last sequence update)  
 DT 15-JUL-1990 (Rel. 38, Last annotation update)  
 DE H-2 CLASS I HISTOCOMPATIBILITY ANTIGEN, D-D ALPHA CHAIN PRECURSOR  
 DE (H-2D(D)).  
 GN H2-D.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 83058139.  
 RA MARGULIES D.H., EVANS G.A., OZATO K., CAMERINI-OTERO R.D., TANAKA K.,  
 RA APPELLA E., SEIDMAN J.G.;  
 RT "Expression of H-2Dd and H-2Id mouse major histocompatibility antigen  
 RT genes in L cells after DNA-mediated gene transfer."  
 RL J. Immunol. 130:463-470(1983).  
 RN [2]  
 RP REVISIONS.  
 RX MEDLINE; 85235604.  
 RA MARGULIES D.H., EVANS G.A., OZATO K., CAMERINI-OTERO R.D., TANAKA K.,  
 RA APPELLA E., SEIDMAN J.G.;  
 RT "Partial nucleotide sequence of H-2Dd major histocompatibility

RT antigen gene.";  
 RL J. Immunol. 135:1537-1537(1985).  
 [3]  
 RP SEQUENCE FROM N.A.  
 RP MEDLINE: 85140250.  
 RX SHER B.T., NAIRN R., COLIGAN J.E., HOOD L.E.;  
 RA "DNA sequence of the mouse H-2Dd transplantation antigen gene.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 82:1175-1179(1985).  
 [4]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=BALB/C;  
 RC MEDLINE: 99086740.  
 RX  
 RA WANG M., STEPKOWSKI S.M., HERBERT J.S., TIAN L., YU J., KAHAN B.D.;  
 RT "Nucleotide sequences of three H-2K and three H-2D complementary DNA  
 RT clones coding mouse class I MHC heavy chain proteins.";  
 RL Ann. Transplant. 1:26-31(1996).  
 [5]  
 RN SEQUENCE OF 25-125.  
 RP MEDLINE: 82046438.  
 RX  
 RA NAIRN R., NATHENSON S.G., COLIGAN J.E.;  
 RT "Amino acid sequence of cyanogen bromide fragment CN-C (residues  
 RT 24-98) of the mouse histocompatibility antigen H-2Dd. A comparison of  
 RT the amino-terminal 100 residues of H-2Dd, Dd, Kd, and Kb reveals  
 RT discrete areas of diversity.";  
 RL Biochemistry 20:4739-4745(1981).  
 [6]  
 RN X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).  
 RP MEDLINE: 98397257.  
 RX  
 RA ACHOUR A., PERSSON K., HARRIS R.A., SUNDBACK J., SENTMAN C.L.,  
 RA LINDQVIST Y., SCHNEIDER G., KARRE K.;  
 RT "The crystal structure of H-2Dd MHC class I complexed with the HIV-1-  
 RT derived peptide P18-I10 at 2.4-A resolution: implications for T cell  
 RT and NK cell recognition.";  
 RL Immunity 9:199-208(1998).  
 CC  
 CC -I- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO  
 CC THE IMMUNE SYSTEM.  
 CC  
 CC -I- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
 CC MICROGLOBULIN).  
 CC  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; L29190; AAA39581.1;  
 DR EMBL; U47326; AAB17604.1;  
 DR PIR; A02201; A02201.  
 DR PDB; 1BII; 14-OCT-98.  
 DR MGD; MGI:95896; H2-D.  
 DR PROSITE; PS00290; IG\_MHC; 1.  
 DR PFAM; PF00047; Ig; 1.  
 DR PFAM; PF00129; MHC\_I; 1.  
 KW MHC I; Transmembrane; Glycoprotein; Signal; 3D-structure.  
 FT SIGNAL 1 24  
 FT CHAIN 25 365 H-2 CLASS I HISTOCOMPATIBILITY ANTIGEN,  
 FT D-D ALPHA CHAIN.  
 FT DOMAIN 25 114  
 FT DOMAIN 115 206 EXTRACELLULAR ALPHA-1.  
 FT DOMAIN 207 298 EXTRACELLULAR ALPHA-2.  
 FT DOMAIN 299 311 EXTRACELLULAR ALPHA-3.  
 FT TRANSMEM 312 334 CONNECTING PEPTIDE.  
 FT DOMAIN 335 365 CYTOPLASMIC TAIL.  
 FT DISULFID 125 188 BY SIMILARITY.  
 FT DISULFID 227 283 BY SIMILARITY.  
 FT CARBOHYD 110 110 POTENTIAL.  
 FT CARBOHYD 200 200 POTENTIAL.  
 SQ SEQUENCE 365 AA; 41110 MW; 91AA5F01 CRC32;

Query Match

76.0%; Score 38; DB 1; Length 365;

Best Local Similarity 80.0%; Pred. No. 3.3;  
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 REDLRTLLRY 10  
 Db 99 RVDLRTALRY 108  
 | | | | | | | |  
 Search completed: February 8, 2000, 01:25:57  
 Job time: 1557 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2000, 19:16:12 ; Search time 176.54 Seconds  
(without alignments)  
3.927 Million cell updates/sec

Title: US-08-653-294-23  
Perfect score: 50  
Sequence: 1 REDLRTLRLY 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 225878 seqs, 69334122 residues

Total number of hits satisfying chosen parameters: 225878

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database :

SPTREMBL\_12:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phase:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50	100.0	39	7 019688	019688 homo sapien
2	50	100.0	90	7 019193	019193 homo sapien
3	50	100.0	181	7 062898	062898 homo sapien
4	50	100.0	181	7 078138	078138 homo sapien
5	50	100.0	181	7 078142	078142 homo sapien
6	50	100.0	274	7 019692	019692 homo sapien
7	50	100.0	322	7 019627	019627 homo sapien
8	50	100.0	359	7 029934	029934 homo sapien
9	50	100.0	362	7 029705	029705 homo sapien
10	50	100.0	362	7 029846	029846 homo sapien
11	50	100.0	362	7 078189	078189 homo sapien
12	45	90.0	137	7 025533	025533 pan troglod
13	45	90.0	330	7 019356	019356 macaca mula
14	45	90.0	331	7 002944	002944 macaca mula
15	45	90.0	331	7 002945	002945 macaca mula
16	45	90.0	346	7 029459	029459 rattus norv
17	45	90.0	346	7 078088	078088 rattus norv
18	45	90.0	348	7 046875	046875 rattus norv
19	44	88.0	362	7 029693	029693 homo sapien
20	44	88.0	362	7 079612	079612 homo sapien

21	43	86.0	98	7	031211	mus musculu
22	43	86.0	362	7	095457	mus musculu
23	40	80.0	89	7	019674	homo sapien
24	40	80.0	89	7	019565	homo sapien
25	40	80.0	131	7	097998	homo sapien
26	40	80.0	131	7	097999	homo sapien
27	40	80.0	181	7	019779	homo sapien
28	40	80.0	181	7	030197	homo sapien
29	40	80.0	181	7	019669	homo sapien
30	40	80.0	181	7	078028	homo sapien
31	40	80.0	355	7	029853	homo sapien
32	40	80.0	361	7	09XXY0	pongo pygma
33	40	80.0	362	7	029637	homo sapien
34	40	80.0	362	7	029935	homo sapien
35	40	80.0	362	7	079524	homo sapien
36	40	80.0	362	7	029850	homo sapien
37	40	80.0	362	7	029661	homo sapien
38	40	80.0	362	7	078180	homo sapien
39	40	80.0	362	7	029933	homo sapien
40	39	78.0	89	4	095956	homo sapien
41	39	78.0	89	7	077959	homo sapien
42	39	78.0	89	7	078077	homo sapien
43	39	78.0	181	7	019555	homo sapien
44	39	78.0	181	7	077933	homo sapien
45	39	78.0	181	7	077935	homo sapien

## ALIGNMENTS

RESULT 1  
019688  
ID 019688 PRELIMINARY; PRT; 39 AA.  
AC 019688;  
DT 01-JAN-1998 (TRENBLrel. 05, Created)  
DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)  
DT 01-NOV-1998 (TRENBLrel. 08, Last annotation update)  
DE HLA-B\*27 VARIANT EXON 2 (ALPHA1 DOMAIN) (FRAGMENT).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA BLASZYK R., WEBER M., SALAMA A.;  
RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.  
DR EMBL; X83727; CAA58698.1; -  
DR PFAM; PF00129; MHC\_I; 1.  
KW MHC.  
FT NON\_TER 1 1  
FT NON\_TER 39 39  
SQ SEQUENCE 39 AA; 4748 MW; 6F714D4C CRC32;

Query Match 100.0%; Score 50; DB 7; Length 39;  
Best Local Similarity 100.0%; Pred. No. 0.0094;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 REDLRTLRLY 10

Db 24 REDLRTLRLY 33

RESULT 2

019193  
ID 019193 PRELIMINARY; PRT; 90 AA.  
AC 019193;  
DT 01-JAN-1998 (TRENBLrel. 05, Created)  
DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)  
DT 01-NOV-1998 (TRENBLrel. 08, Last annotation update)  
DE HISTOCOMPATIBILITY ANTIGEN ALPHA 1 DOMAIN (FRAGMENT).  
GN HLA-B\*27.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 92337445.  
 RA HIGGINS C.M., LUND T., SHIPLEY M.E., EBRINGER A.,  
 RA SADOWSKA-WROBLEWSKA M., CRAIG R.K.;  
 RT "Ankylosing spondylitis and HLA-B\*27: restriction fragment length  
 RT polymorphism and sequencing of an HLA-B\*27 allele from a patient with  
 RT ankylosing spondylitis";  
 RL Ann. Rheum. Dis. 51:855-862(1992).  
 DR EMBL: S39758; CAB27364.1; -;  
 DR PFAM: PF00129; MHC\_I; 1.  
 KW MHC.  
 FT NON\_TER 1 1  
 FT NON\_TER 90 90  
 SQ SEQUENCE 90 AA; 10571 MW; F22CCB4E CRC32;

Query Match 100.0%; Score 50; DB 7; Length 90;  
 Best Local Similarity 100.0%; Pred. No. 0.022;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 REDLRTLLRY 10  
 Db 75 REDLRTLLRY 84

RESULT 3  
 O62898 PRELIMINARY; PRT; 181 AA.  
 ID O62898  
 AC O62898;  
 DT 01-AUG-1998 (TEMBLrel. 07, Created)  
 DT 01-AUG-1998 (TEMBLrel. 07, Last sequence update)  
 DT 01-NOV-1999 (TEMBLrel. 12, Last annotation update)  
 DE MHC CLASS I ANTIGEN (FRAGMENT).  
 GN HLA-B.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA KOSMAN C.A., HURLEY C.K.;  
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF054012; AAC32563.1; -;  
 DR EMBL: AF054011; AAC32563.1; JOINED.  
 DR HSP: P10318; 1ROG.  
 DR PFAM: PF00129; MHC\_I; 1.  
 KW MHC.  
 FT NON\_TER 1 1  
 FT NON\_TER 181 181  
 SQ SEQUENCE 181 AA; 21107 MW; D8E533DD CRC32;

Query Match 100.0%; Score 50; DB 7; Length 181;  
 Best Local Similarity 100.0%; Pred. No. 0.045;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 REDLRTLLRY 10  
 Db 74 REDLRTLLRY 83

RESULT 4  
 O78138 PRELIMINARY; PRT; 181 AA.  
 ID O78138  
 AC O78138;  
 DT 01-NOV-1998 (TEMBLrel. 08, Created)  
 DT 01-NOV-1998 (TEMBLrel. 08, Last sequence update)  
 DT 01-NOV-1999 (TEMBLrel. 12, Last annotation update)  
 DE MHC CLASS I ANTIGEN (FRAGMENT).  
 GN HLA-B.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.

OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA KOSMAN C.A., HURLEY C.K.;  
 RT "Novel HLA Class I B Locus Alleles";  
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF071770; AAC35939.1; -;  
 DR EMBL: AF071769; AAC35939.1; JOINED.  
 DR HSP: P10318; 1ROG.  
 DR PFAM: PF00129; MHC\_I; 1.  
 KW MHC.  
 FT NON\_TER 1 1  
 FT NON\_TER 181 181  
 SQ SEQUENCE 181 AA; 21103 MW; 8CF468CF CRC32;

Query Match 100.0%; Score 50; DB 7; Length 181;  
 Best Local Similarity 100.0%; Pred. No. 0.045;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 REDLRTLLRY 10  
 Db 74 REDLRTLLRY 83

RESULT 5  
 O78142 PRELIMINARY; PRT; 181 AA.  
 ID O78142  
 AC O78142;  
 DT 01-NOV-1998 (TEMBLrel. 08, Created)  
 DT 01-NOV-1998 (TEMBLrel. 08, Last sequence update)  
 DT 01-NOV-1999 (TEMBLrel. 12, Last annotation update)  
 DE MHC CLASS I ANTIGEN (FRAGMENT).  
 GN HLA-B.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STEINER N.K., HURLEY C.K., KOESTER R.P.;  
 RT "Novel-HLA-B allele";  
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF072764; AAC25779.1; -;  
 DR EMBL: AF072763; AAC25779.1; JOINED.  
 DR HSP: P10318; 1ROG.  
 DR PFAM: PF00129; MHC\_I; 1.  
 KW MHC.  
 FT NON\_TER 1 1  
 FT NON\_TER 181 181  
 SQ SEQUENCE 181 AA; 21079 MW; 24949B0F CRC32;

Query Match 100.0%; Score 50; DB 7; Length 181;  
 Best Local Similarity 100.0%; Pred. No. 0.045;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 REDLRTLLRY 10  
 Db 74 REDLRTLLRY 83

RESULT 6  
 O19692 PRELIMINARY; PRT; 274 AA.  
 ID O19692  
 AC O19692;  
 DT 01-JAN-1998 (TEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TEMBLrel. 05, Last sequence update)  
 DT 01-NOV-1999 (TEMBLrel. 12, Last annotation update)  
 DE MHC CLASS I HLA-B\*27 M (FRAGMENT).  
 GN HLA-B.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

RN RP SEQUENCE FROM N.A.
RX MEDLINE: 87009855.
RA COPPIN H.L., MCDEVITT H.O.;
RT "Absence of polymorphism between HLA-B27 genomic exon sequences
RT isolated from normal donors and ankylosing spondylitis patients.";
RL J. Immunol. 137:2168-2172(1986).
DR EMBL; M14013; AAA59643.1; -.
DR HSSP; P10318; IROG.
DR PROSITE; PS00290; IG_MHC; 1.
DR PFAM; PF00047; Ig; 1.
DR PFAM; PF00129; MHC_I; 1.
KW MHC.
FT NON_TER 1
FT NON_TER 274
SQ SEQUENCE 274 AA; 31659 MW; 9A74A6BA CRC32;

Query Match 100.0%; Score 50; DB 7; Length 274;
Best Local Similarity 100.0%; Pred. No. 0.07;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 REDRTLRLY 10
Db 75 REDRTLRLY 84

RESULT 7
ID 019627 PRELIMINARY; PRT; 322 AA.
AC 019627;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
DE HLA-B37 (FRAGMENT).
GN B-3701.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN RP SEQUENCE FROM N.A.
RA HURLEY C.K., BEI M., RODRIGUEZ S., JOHNSON A.;
RL Submitted (JUN-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL; U11267; AAA19927.1; -.
DR HSSP; P30685; IA9E.
DR PROSITE; PS00290; IG_MHC; 1.
DR PFAM; PF00047; Ig; 1.
DR PFAM; PF00129; MHC_I; 1.
KW MHC.
FT NON_TER 322
FT NON_TER 322
SQ SEQUENCE 322 AA; 36626 MW; DF3B7744 CRC32;

Query Match 100.0%; Score 50; DB 7; Length 322;
Best Local Similarity 100.0%; Pred. No. 0.082;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 REDRTLRLY 10
Db 99 REDRTLRLY 108

RESULT 8
ID 029934 PRELIMINARY; PRT; 359 AA.
AC 029934;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
DE HLA-B27 (FRAGMENT).
GN HLA-B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

```

```

OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN RP SEQUENCE FROM N.A.
RX MEDLINE: 86149317.
RA SZOTS H., RITHMULLER G., WEISS E., MEO T.;
RT "Complete sequence of HLA-B27 cDNA identified through the
RT characterization of structural markers unique to the HLA-A, -B, and -C
RT allelic series.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:1428-1432(1986).
DR EMBL; M12678; AAA59614.1; -.
DR HSSP; P10318; IROG.
DR PROSITE; PS00290; IG_MHC; 1.
DR PFAM; PF00047; Ig; 1.
DR PFAM; PF00129; MHC_I; 1.
KW MHC.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 359 AA; 40042 MW; 069F7E64 CRC32;

Query Match 100.0%; Score 50; DB 7; Length 359;
Best Local Similarity 100.0%; Pred. No. 0.092;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 REDRTLRLY 10
Db 96 REDRTLRLY 105

RESULT 9
ID 029705 PRELIMINARY; PRT; 362 AA.
AC 029705;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
DE MHC CLASS I ANTIGEN HLA-B PRECURSOR.
GN HLA-B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN RP SEQUENCE FROM N.A.
RA BALAS A., SANTOS S., VICARIO J.L.;
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U31971; AAA98506.1; -.
DR HSSP; P10318; IROG.
DR PROSITE; PS00290; IG_MHC; 1.
DR PFAM; PF00047; Ig; 1.
DR PFAM; PF00129; MHC_I; 1.
KW SIGNAL; MHC.
FT SIGNAL 1
FT SIGNAL 24
FT CHAIN 25
FT CHAIN 362
SQ SEQUENCE 362 AA; 40479 MW; 09C9D20A CRC32;

Query Match 100.0%; Score 50; DB 7; Length 362;
Best Local Similarity 100.0%; Pred. No. 0.093;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 REDRTLRLY 10
Db 99 REDRTLRLY 108

RESULT 10
ID 029846 PRELIMINARY; PRT; 362 AA.
AC 029846;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
DE HUMAN LYMPHOCYTE ANTIGEN HLA-B27.

```

OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 [1]  
 RN RP SEQUENCE FROM N.A.  
 RC TISSUE=BLOOD;  
 RX MEDLINE: 94375872.  
 RA DEL PORTO P., D'AMATO M., FIORILLO M.T., TUOSTO L., PICCOLELLA E.,  
 RA SORRENTINO R.;  
 RT "Identification of a novel HLA-B\*27 subtype by restriction analysis of  
 RT a cytotoxic gamma delta T cell clone.";  
 RL J. Immunol. 153:3093-3100(1994).  
 DR EMBL: Z33453; CAA83876.1; -  
 DR HSSP: P10318; IROG  
 DR PROSITE: PS00290; IG\_MHC; 1.  
 DR PFAM: PF00047; Ig; 1.  
 DR PFAM: PF00129; MHC\_I; 1.  
 KW MHC.  
 SQ SEQUENCE 362 AA; 40450 MW; CCA23A50 CRC32;

Query Match 100.0%; Score 50; DB 7; Length 362;  
 Best Local Similarity 100.0%; Pred. No. 0.093;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 REDRLTLRY 10  
 |||||  
 Db 99 REDRLTLRY 108

RESULT 11  
 O78189 PRELIMINARY; PRT; 362 AA.  
 AC O78189;  
 DT 01-NOV-1998 (TEMBLrel. 08, Created)  
 DT 01-NOV-1998 (TEMBLrel. 08, Last sequence update)  
 DT 01-NOV-1999 (TEMBLrel. 12, Last annotation update)  
 DE MHC CLASS I ANTIGEN.  
 GN HLA-B.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 [1]  
 RN RP SEQUENCE FROM N.A.  
 RA SEURYNCK K.L., BAXTER-LOWE L.A.;  
 RT "B27052 w496D.";  
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF026218; AAC42275.1; -  
 DR HSSP: P10318; IROG.  
 DR PROSITE: PS00290; IG\_MHC; 1.  
 DR PFAM: PF00047; Ig; 1.  
 DR PFAM: PF00129; MHC\_I; 1.  
 KW MHC.  
 SQ SEQUENCE 362 AA; 40486 MW; 2B0EF602 CRC32;

Query Match 100.0%; Score 50; DB 7; Length 362;  
 Best Local Similarity 100.0%; Pred. No. 0.093;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 REDRLTLRY 10  
 |||||  
 Db 99 REDRLTLRY 108

RESULT 12  
 Q95533 PRELIMINARY; PRT; 137 AA.  
 AC Q95533;  
 DT 01-FEB-1997 (TEMBLrel. 02, Created)  
 DT 01-FEB-1997 (TEMBLrel. 02, Last sequence update)  
 DT 01-NOV-1999 (TEMBLrel. 08, Last annotation update)  
 DE CLASS I HISTOCOMPATIBILITY ANTIGEN (FRAGMENT).

GN HLA-B.  
 OS Pan troglodytes (Chimpanzee).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Pan.  
 [1]  
 RN RP SEQUENCE FROM N.A.  
 RC STRAIN=WODKA;  
 RX MEDLINE: 94286544.  
 RA MCADAM S.N., BOYSON J.E., LIU X., GARBER T.L., HUGHES A.L.,  
 RA BONTROP R.E., WATKINS D.I.;  
 RT "A uniquely high level of recombination at the HLA-B locus.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 91:5893-5897(1994).  
 DR EMBL: U05585; AAA50188.1; -  
 DR PFAM: PF00129; MHC\_I; 1.  
 KW MHC.  
 FT NON\_TER 1  
 FT NON\_TER 137  
 SQ SEQUENCE 137 AA; 15922 MW; B316D3BC CRC32;

Query Match 90.0%; Score 45; DB 7; Length 137;  
 Best Local Similarity 90.0%; Pred. No. 0.3;  
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 REDRLTLRY 10  
 ||:|||||  
 Db 40 RENRLTLRY 49

RESULT 13  
 O19356 PRELIMINARY; PRT; 330 AA.  
 AC O19356;  
 DT 01-JAN-1998 (TEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TEMBLrel. 05, Last sequence update)  
 DT 01-NOV-1999 (TEMBLrel. 12, Last annotation update)  
 DE MHC CLASS I ANTIGEN (FRAGMENT).  
 GN MAMU-AG.  
 OS Macaca mulatta (Rhesus macaque).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Cercopitheidae; Cercopitheinae;  
 OC Macaca.  
 [1]  
 RN RP SEQUENCE FROM N.A.  
 RC TISSUE=PLACENTA;  
 RA BOYSON J.E., IWANAGA K.K., GOLOS T.G., WATKINS D.I.;  
 RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: U84785; AAB58547.1; -  
 DR HSSP: P30491; IALM.  
 DR PROSITE: PS00290; IG\_MHC; 1.  
 DR PFAM: PF00047; Ig; 1.  
 DR PFAM: PF00129; MHC\_I; 1.  
 KW MHC.  
 FT NON\_TER 1  
 FT NON\_TER 330  
 SQ SEQUENCE 330 AA; 37654 MW; 7403A522 CRC32;

Query Match 90.0%; Score 45; DB 7; Length 330;  
 Best Local Similarity 90.0%; Pred. No. 0.73;  
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 REDRLTLRY 10  
 ||:|||||  
 Db 90 RENRLTLRY 99

RESULT 14  
 O02944 PRELIMINARY; PRT; 331 AA.  
 AC O02944;  
 DT 01-JUL-1997 (TEMBLrel. 04, Created)  
 DT 01-JUL-1997 (TEMBLrel. 04, Last sequence update)  
 DT 01-NOV-1999 (TEMBLrel. 12, Last annotation update)



DE MHC CLASS I ANTIGEN (FRAGMENT).  
GN MAMU-AG.  
OS Macaca mulatta (Rhesus macaque).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecinae;  
OC Macaca.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-AMNIOTIC MEMBRANE;  
RA BOYSON J.E., IWANAGA K.K., GOLOS T.G., WATKINS D.I.;  
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U84784; AAB58346.1; -;  
DR HSSP; P30491; IALM.  
DR PROSITE; PS00290; IG\_MHC; 1.  
DR PFAM; PF00047; ig; 1.  
DR PFAM; PF00129; MHC\_I; 1.  
KW MHC.  
FT NON\_TER 1  
SQ SEQUENCE 331 AA; 37767 MW; 17A6473C CRC32;

Query Match 90.0%; Score 45; DB 7; Length 331;  
Best Local Similarity 90.0%; Pred. No. 0.73;  
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 REDRLTLRY 10  
II:|||||  
DB 91 RENLRLTRY 100

RESULT 15  
O02945  
ID O02945 PRELIMINARY; PRT; 331 AA.  
AC O02945;  
DT 01-JUL-1997 (TRENBLrel. 04, Created)  
DT 01-JUL-1997 (TRENBLrel. 04, Last sequence update)  
DT 01-NOV-1999 (TRENBLrel. 12, Last annotation update)  
DE MHC CLASS I ANTIGEN (FRAGMENT).  
GN MAMU-AG.  
OS Macaca mulatta (Rhesus macaque).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecinae;  
OC Macaca.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA BOYSON J.E., IWANAGA K.K., GOLOS T.G., WATKINS D.I.;  
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U84786; AAB58348.1; -;  
DR HSSP; P30491; IALM.  
DR PROSITE; PS00290; IG\_MHC; 1.  
DR PFAM; PF00047; ig; 1.  
DR PFAM; PF00129; MHC\_I; 1.  
KW MHC.  
FT NON\_TER 1  
SQ SEQUENCE 331 AA; 37800 MW; D4E0FF2E CRC32;

Query Match 90.0%; Score 45; DB 7; Length 331;  
Best Local Similarity 90.0%; Pred. No. 0.73;  
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 REDRLTLRY 10  
II:|||||  
DB 91 RENLRLTRY 100

Search completed: February 8, 2000, 19:16:13  
Job time: 21500 sec

**THIS PAGE BLANK (USPTO)**

OM of: US-08-653-294-23 to: GenEmbl.\* out\_format : pfs

Date: Feb 8, 2000 10:22 PM

About: Results were produced by the GenCore software, version 4.5.  
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp  
-O=/cgl1\_1/USPTO\_spool/US08653294/runat\_04022000\_160701\_15779/app\_query.fasta.2  
-DB=GenEmbl -QPMI=fastp -SUFFIX=rge -GAPOP=12.000 -GAPEXT=4.500  
-MINMATCH=0.100 -LOOPEXT=0.000 -LOOPEXT=0.000 -GAPOP=4.500  
-GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000  
-FGAEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000  
-DELEXT=7.000 -START=1 -MATRIX=blosom62 -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR\_SCORE=pct -ALIGN=15 -MODE=LOCAL  
-OUTFM=pfs -NORM=ext -MINLEN=0 -MAXLEN=1000000 -USER=US08653294  
-NCPU=6 -ICPU=3 -NO\_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-08-653-294-23

Query length: 10

Database: GenEmbl.\*

Database sequences: 821193

Database length: -1518192014

Search time (sec): 10176.920000

score\_list:

Sequence	Strtd Orig	ZScore	Escore	Len	Documentation
gb_pr1:HSAB27V2	50.00	190.46	0.0138	195	XG3727 H.sapiens HLA-B27 variant
gb_pr3:HSAB27HLA1	50.00	187.73	0.0196	270	AF072763 Homo sapiens MHC class
gb_pr3:HSB1524V1	50.00	187.73	0.0196	270	AF054011 Homo sapiens isolate G
gb_pr3:HSBCHL8C1	50.00	187.73	0.0196	270	AF071769 Homo sapiens MHC class
gb_pr4:HS27052B1	50.00	187.73	0.0196	270	AF102563 Homo sapiens MHC class
gb_pr4:HSBHLAWBY2	50.00	187.73	0.0196	270	AF110257 Homo sapiens MHC class
gb_pr2:HSBHLAWBY2	50.00	181.82	0.0418	546	L76095 Homo sapiens MHC class I
gb_pr1:HSBHLAWBY2	50.00	178.39	0.0650	822	M14013 Human MHC class I HLA-B2
gb_pr1:HSU11267	50.00	177.01	0.0775	968	U11267 Human HLA-B37 (B-3701) m
gb_pr1:HSU11267	50.00	176.60	0.0818	1017	M62852 Human MHC class I HLA-E
gb_pr1:HSU11267	50.00	176.53	0.0825	1026	A28264 H.sapiens mRNA for HLA-E
gb_pr1:HSU11267	50.00	176.53	0.0825	1026	E01342 cDNA encoding C-termina
gb_pr1:HSU11267	50.00	176.06	0.0876	1084	L20086 Human MHC class I (HLA-
gb_pr2:HSU31971	50.00	176.03	0.0880	1089	U31971 Human MHC class I antiq
gb_pr3:AF026218	50.00	176.03	0.0880	1089	AF026218 Homo sapiens MHC clas
gb_pr1:HSUHLA8A4	50.00	176.00	0.0884	1093	M19756 Human MHC class I HLA-E
gb_pr1:HSUHLA8A4	50.00	176.00	0.0884	1093	M32320 Human MHC HLA protein,
gb_pr1:HSUHLA8A4	50.00	174.93	0.1013	1241	X03665 Human class I MHC gene
gb_pr1:HSUHLA8A4	50.00	174.72	0.1040	1272	Z33453 H.sapiens mRNA for huma
gb_pr1:HSUHLA8A4	50.00	173.30	0.1249	1507	M12678 Human HLA-B27 mRNA, com
gb_pr1:HSUHLA8A4	50.00	165.88	0.3235	3649	M54883 Human MHC class I HLA-E
gb_pr1:HSUHLA8A4	50.00	165.07	0.3586	4015	E01341 Genomic DNA encoding HI
gb_pr1:HSUHLA8A4	50.00	164.61	0.3804	4016	M12967 Human MHC class I HLA-E
gb_pr1:HSUHLA8A4	50.00	164.52	0.3848	4242	S39758 HLA-B27 (HLA-B*2705)-h
gb_pr1:HSUHLA8A4	50.00	160.96	0.6076	6553	X03945 Human gene for HLA-B27
gb_pr1:HSUHLA8A4	45.00	164.26	0.3981	412	U05585 Pan troglodytes class I
gb_pr1:HSUHLA8A4	45.00	156.48	1.108	1041	AJ004889 Rattus norvegicus mRNA
gb_pr1:HSUHLA8A4	45.00	156.40	1.109	1050	R08530 R.norvegicus mRNA for M
gb_pr1:HSUHLA8A4	45.00	156.24	1.111	1071	Y13890 Rattus norvegicus mRNA
gb_pr3:MMU41831	45.00	156.15	1.113	1082	U41831 Macaca mulatta MHC clas
gb_pr2:MMU84786	45.00	153.93	1.50	1410	U84786 Macaca mulatta MHC clas
gb_pr2:MMU84786	45.00	153.90	1.50	1415	U84784 Macaca mulatta MHC clas
gb_pr2:MMU84786	45.00	153.72	1.54	1445	U84785 Macaca mulatta MHC clas
gb_pr2:MMU84786	44.00	163.82	0.4211	1089	AF107578 Homo sapiens MHC class
gb_pr4:HSBCHL8C1	44.00	152.11	1.89	1089	X73578 H. sapiens HLA-B*2706 m
gb_pr1:HSBHLA8A4	44.00	152.11	1.89	1089	U27608 Human MHC class I antiq
gb_pr3:HSU35734	44.00	152.11	1.89	1089	U35734 Human MHC class I antiq
gb_pr1:HSB3043	44.00	152.03	1.91	1100	D83043 Human HLA-B mRNA, allel
gb_pr1:HSB3043	44.00	114.27	242.30	98836	AL132774 Homo sapiens chromos
gb_pr1:HSB3043	43.00	159.09	0.7723	295	J00409 Mouse MHC class I H2 gen
gb_pr1:HSB3043	43.00	149.92	2.51	880	M58595 Mouse MHC class I H-2D-M

gb\_ro:MUSMHHKQO + 43.00 145.81 4.24 1435 ! M14827 Mouse MHC class I H-  
gb\_ro:MMU47327 + 43.00 145.07 4.66 1567 ! U47327 Mus musculus MHC cla  
gb\_ro:MUSMHH2DK + 43.00 135.84 15.24 4707 ! M18524 Mouse MHC class I H-

seq\_name: gb\_pr1:HSAB27V2

seq\_documentation\_block:

LOCUS HSLAB27V2 195 bp DNA PRI 31-JAN-1995  
DEFINITION H.sapiens HLA-B27 variant gene (exon 2).  
ACCESSION X83727  
VERSION X83727.1 GI:663002  
KEYWORDS HLA-B gene; human leukocyte antigen; major histocompatibility complex class I.  
SOURCE human.

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;  
Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 195)  
AUTHORS Blasczyk,R., Weber,M. and Salama,A.  
JOURNAL Unpublished

REFERENCE 2 (bases 1 to 195)  
AUTHORS Blasczyk,R.  
JOURNAL Direct Submission

TITLE Submitted (08-JAN-1995) R. Blasczyk, Bloodbank, Dept.of Intern.  
Medicine, Div of Hematol. and Oncolog., Spandauer Damm 130, Univ.  
Hosp. Rudolf Virchow, Freie Univ., D-14050 Berlin, FRG  
JOURNAL

FEATURES  
Location/Qualifiers

1..195

/organism="Homo sapiens"

/isolate="1235"

/db\_xref="taxon:9606"

/chromosome="6"

/cell\_type="EBV transformed cell line"

/map="p21.3"

<76..>195

/codon\_start=3

/product="HLA-B27 variant exon 2 (alpha domain)"

/protein\_id="CAA58698.1"

/db\_xref="GI:663003"

/db\_xref="SPTREMBL:O19688"

/translation="IEQEGPEYWDRETOICKAKAQTDRDLTLRLYYNQSEA"

45 a 55 c 73 g 22 t

BASE COUNT

ORIGIN

alignment\_scores:

Quality: 50.00 Length: 10

Ratio: 5.000 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-08-653-294-23 x HSLAB27V2 ..

Align seg 1/1 to: HSLAB27V2 from: 1 to: 195

1 ArgGluspleuArgThrLeuLeuArgTyr 10

|||||

147 CGAGAGGACCTGGGACCTGCTCCGCTAC 176

seq\_name: gb\_pr3:HSAB27HLA1

seq\_documentation\_block:

LOCUS HSAB27HLA1 270 bp DNA PRI 08-JUL-1998  
DEFINITION Homo sapiens MHC class I antigen HLA-B gene (HLA-B\*27 variant allele), exon 2.

ACCESSION AF072763

VERSION AF072763.1 GI:3293562

KEYWORDS

1 of 2

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 270)  
 AUTHORS Steiner,N.K., Hurley,C.K. and Koester,R.P.  
 TITLE Novel-HLA-B allele  
 JOURNAL Unpublished

REFERENCE 2 (bases 1 to 270)  
 AUTHORS Steiner,N.K., Hurley,C.K. and Koester,R.P.  
 TITLE Direct Submission  
 JOURNAL Submitted (21-JUN-1998) Microbiology and Immunology, Georgetown University Medical Center, 3970 Reservoir Road NW, Washington, DC 20007, USA

## FEATURES

source Location/Qualifiers  
 1..270  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 exon 1..270  
 /gene="HLA-B"  
 /number=2  
 BASE COUNT 54 a 87 c 93 g 36 t  
 ORIGIN

alignment\_scores:  
 Quality: 50.00 Length: 10  
 Ratio: 5.000 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-08-653-294-23 x HSB27HLA1 ..

Align seg 1/1 to: HSB27HLA1 from: 1 to: 270

1 ArgGluAspLeuArgThrLeuLeuArgTyr 10  
 |||||  
 222 CGAGAGACCTGCGGACCTGCTCGCTAC 251

seq\_name: gb\_pr3:HSB1524V1

seq\_documentation\_block: 270 bp DNA PRI 22-AUG-1998  
 LOCUS HSB1524V1  
 DEFINITION Homo sapiens isolate GN00211 MHC class I antigen HLA-B gene (B\*1543 allele), exon 2.

ACCESSION AF054011  
 VERSION AF054011.1 GI:2984766  
 KEYWORDS  
 SEGMENT 1 of 2  
 SOURCE human.

## ORGANISM

Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 270)  
 AUTHORS Kosman,C.A. and Hurley,C.K.  
 TITLE Novel HLA Class I B locus alleles  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 270)  
 AUTHORS Kosman,C.A. and Hurley,C.K.  
 TITLE Direct Submission  
 JOURNAL Submitted (18-MAR-1998) Microbiology & Immunology, Georgetown University, 3970 Reservoir Rd. NW, Washington, DC 20007, USA

## FEATURES

source Location/Qualifiers  
 1..270  
 /organism="Homo sapiens"  
 /isolate="GN00211"  
 /db\_xref="taxon:9606"  
 exon 1..270  
 /gene="HLA-B"  
 /number=2  
 BASE COUNT 57 a 90 c 86 g 37 t  
 ORIGIN

alignment\_scores:  
 Quality: 50.00 Length: 10  
 Ratio: 5.000 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-08-653-294-23 x HSB1524V1 ..

Align seg 1/1 to: HSB1524V1 from: 1 to: 270

1 ArgGluAspLeuArgThrLeuLeuArgTyr 10  
 |||||  
 222 CGAGAGACCTGCGGACCTGCTCGCTAC 251

seq\_name: gb\_pr3:HSMHCHLBC1

## seq\_documentation\_block:

LOCUS HSMCHLBC1 270 bp DNA PRI 17-SEP-1998  
 DEFINITION Homo sapiens MHC class I antigen HLA-B gene (HLA-B\*5303 allele), exon 2.

ACCESSION AF071769  
 VERSION AF071769.1 GI:3243269  
 KEYWORDS  
 SEGMENT 1 of 2  
 SOURCE human.

## ORGANISM

Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 270)

AUTHORS Kosman,C.A. and Hurley,C.K.

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 270)

AUTHORS Kosman,C.A. and Hurley,C.K.

JOURNAL Direct Submission

TITLE Submitted (12-JUN-1998) Microbiology & Immunology, Georgetown University, 3970 Reservoir Rd. N.W., Washington, DC 20007, USA

## FEATURES

source Location/Qualifiers  
 1..270  
 /organism="Homo sapiens"  
 /isolate="GN00231"  
 /db\_xref="taxon:9606"  
 exon 1..270  
 /gene="HLA-B"  
 /number=2  
 BASE COUNT 59 a 90 c 84 g 37 t  
 ORIGIN

## alignment\_scores:

Quality: 50.00 Length: 10  
 Ratio: 5.000 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-08-653-294-23 x HSMCHLBC1 ..

Align seg 1/1 to: HSMCHLBC1 from: 1 to: 270

1 ArgGluAspLeuArgThrLeuLeuArgTyr 10  
 |||||  
 222 CGAGAGACCTGCGGACCTGCTCGCTAC 251

seq\_name: gb\_pr4:HS27052B1

## seq\_documentation\_block:

LOCUS HS27052B1 270 bp DNA PRI 21-JUN-1999  
 DEFINITION Homo sapiens MHC class I antigen HLA-B gene, HLA-B\*2716 allele, exon 2.

ACCESSION AF102563  
 VERSION AF102563.1 GI:4704574  
 KEYWORDS  
 SEGMENT 1 of 2  
 SOURCE human.

## ORGANISM

Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 270)  
AUTHORS Kosman,C.A. and Hurley,C.K.  
TITLE Novel Class I HLA-B Alleles  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 270)  
AUTHORS Kosman,C.A. and Hurley,C.K.  
TITLE Direct Submission  
JOURNAL Submitted (28-OCT-1998) Microbiology and Immunology, Georgetown University, 3970 Reservoir Rd. N.W., Washington, DC 20007, USA  
FEATURES  
source Location/Qualifiers  
1..270  
/organism="Homo sapiens"  
/isolate="GN00246"  
/db\_xref="taxon:9606"  
1..270  
/gene="HLA-B"  
/number=2

BASE COUNT 56 a 88 c 90 g 36 t  
ORIGIN

alignment\_scores:  
Quality: 50.00 Length: 10  
Ratio: 5.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-08-653-294-23 x HS27052B1 ..

Align seg 1/1 to: HS27052B1 from: 1 to: 270

1 ArgGluAspLeuArgThrLeuLeuArgTyr 10  
|||||  
222 CGAGAGAGACTGCGGACCTGCTCCGCTAC 251

seq\_name: gb\_pr4: HSHLAWMBY2

seq\_documentation\_block:  
LOCUS HSHLAWMBY2 270 bp DNA PRI 06-APR-1999  
DEFINITION Homo sapiens MHC class I antigen HLA-B gene (HLA-B\*27 allele), exon 2.

ACCESSION AF110257  
VERSION AF110257.1 GI:4566542

KEYWORDS

SEGMENT 2 of 3

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 270)

Wu,J., Bassinger,S., Montoya,G.D., Yee,J., Griffith,B.B.,

Kearns,J., McKeen,M., Birkos,S., Kamoun,M. and Williams,T.M.

Identification of new HLA-B alleles in potential bone marrow donors

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 270)

Wu,J., Bassinger,S., Montoya,G.D., Yee,J., Griffith,B.B.,

Kearns,J., McKeen,M., Birkos,S., Kamoun,M. and Williams,T.M.

Direct Submission

JOURNAL Submitted (30-NOV-1998) Pathology, Univ. New Mexico, 915 Camino de

Salud, NE, Albuquerque, NM 87131, USA

FEATURES Location/Qualifiers

source 1..270

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/map="6p23"

/chromosome="6"

1..270

/gene="HLA-B"

/number=2

BASE COUNT 54 a 87 c 93 g 36 t

ORIGIN

alignment\_scores:  
Quality: 50.00 Length: 10  
Ratio: 5.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-08-653-294-23 x HSHLAWMBY2 ..

Align seg 1/1 to: HSHLAWMBY2 from: 1 to: 270

1 ArgGluAspLeuArgThrLeuLeuArgTyr 10  
|||||  
222 CGAGAGAGACTGCGGACCTGCTCCGCTAC 251

seq\_name: gb\_pr2: HUMB27052G

seq\_documentation\_block:

LOCUS HUMB27052G 546 bp DNA PRI 27-FEB-1996  
DEFINITION Homo sapiens MHC class I HLA-B\*27052 gene, exons.

ACCESSION L76095

VERSION L76095.1 GI:1203957

KEYWORDS cell surface antigen; cell surface glycoprotein; class I gene;  
integral membrane protein; major histocompatibility complex.

SOURCE Homo sapiens DNA.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 546)

AUTHORS Marcos,C.Y., Fernandez-Vina,M.A., Lazaro,A.M. and Stastny,P.

TITLE Novel HLA-B Alleles

JOURNAL Unpublished (1996)

FEATURES Location/Qualifiers

source 1..546

/organism="Homo sapiens"

/db\_xref="taxon:9606"

1..546

/gene="B-2705v"

1..270

/gene="B-2705v"

271..546

/gene="B-2705v"

454

/gene="B\*27052"

/replace="t"

454

/gene="B\*27052"

109 a 172 c 195 g 70 t

BASE COUNT

ORIGIN

alignment\_scores:

Quality: 50.00 Length: 10

Ratio: 5.000 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-08-653-294-23 x HUMB27052G ..

Align seg 1/1 to: HUMB27052G from: 1 to: 546

1 ArgGluAspLeuArgThrLeuLeuArgTyr 10  
|||||  
222 CGAGAGAGACTGCGGACCTGCTCCGCTAC 251

seq\_name: gb\_pr1: HUMMB2

seq\_documentation\_block:

LOCUS HUMMB2 822 bp DNA PRI 07-JAN-1995  
DEFINITION Human MHC class I HLA-B\*27 M+ gene, exons 2-4 (introns unsequenced).

ACCESSION M14013

VERSION M14013.1 GI:187743

KEYWORDS

```

SOURCE      Human DNA.
ORGANISM    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 822)
AUTHORS    Coppin,H.L. and McDewitt,H.O.
TITLE      Absence of polymorphism between HLA-B*27 genomic exon sequences
            isolated from normal donors and ankylosing spondylitis patients
JOURNAL    J. Immunol. 137 (7), 2168-2172 (1986)
MEDLINE    87009855
FEATURES    Location/Qualifiers
            source
            1..822
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /map="6p21.3"
            /genes="HLA-B"
            /notes="HLA-B"
            /codon_start=1
            /db_xref="GDB:G00-120-048"
            /protein_id="AAA59643.1"
            /db_xref="GI:187744"
            /translation="GSHSMRYFHTSVSRGGRPEFTVGYVDDTLFVRFSDAASPR
            EPRAPWIEQEGVDRETQICKAQTDREDLRLRYNQSEAGSHTLQNMVYCD
            VPDGCLRGYHQDAVDKDYIALNEDLSSTWADPTAAQITQKWEAAVAGQLRAYL
            EGCEVWLRYLENGKETLQRAADPPKXTHVTHPISDHEATLRCWALGFPYPAITLWQ
            RGEQDTQTELVTETPAGDRTFKWAAVVPFSGEQRVYCHVQHEGLPKPLTLRW"
BASE COUNT 176 a 256 c 277 g 113 t
ORIGIN

alignment_scores:
  Quality: 50.00      Length: 10
  Ratio: 5.000       Gaps: 0
  Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-23 x HUMHBM2 ..
Align seg 1/1 to: HUMHBM2 from: 1 to: 822
1 ArgGluAspLeuArgThrLeuLeuArgTyr 10
|||||
223 CGAGAGGACCTGCGGACCTGCTCGCTAC 252

seq_name: gb_prl:HSU11267

seq_documentation_block:
LOCUS      HSU11267 968 bp mRNA PRI 21-JUL-1994
DEFINITION Human HLA-B*37 (B-3701) mRNA, partial cds.
ACCESSION  U11267
VERSION     U11267.1 GI:511785
KEYWORDS
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (sites)
AUTHORS    Hurley,C.K., Bel,M., Rodriguez,S. and Johnson,A.
TITLE      HLA-B*71
JOURNAL    Unpublished
REFERENCE   2 (bases 1 to 968)
AUTHORS    Hurley,C.K.
TITLE      Direct Submission
JOURNAL    Submitted (23-JUN-1994) Carolyn K. Hurley, Microbiology, Georgetown
            University School of Medicine, 3900 Reservoir Road, N.W.,
            Washington, D.C. 20007 USA
FEATURES    Location/Qualifiers
            source
            1..968
            /organism="Homo sapiens"
            /isolate="GU2760"

```

```

            /db_xref="taxon:9606"
            /haplotype="HLA-A2, CW6, B37 / A-, CW7, B71"
            /cell_line="B lymphoblastoid"
            /chromosome="6"
            1..968
            /gene="B-3701"
            1..>968
            /gene="B-3701"
            /citation=[1]
            /codon_start=1
            /function="peptide presentation; histocompatibility
            molecule"
            /product="HLA-B*37"
            /protein_id="AAA19927.1"
            /db_xref="GI:511786"
            /translation="MRVTAPRTLLLLKCAVALTETWAGSHSMRYFHTSVSRGGRGE
            RFTSVGYVDDTQFVRFSDAASPRTEPRAPWIEQEGPEYWDRETOISKTQTYREDL
            RTLRYNQSEAGSHTIQRMSCDVGPRGLRGYNQFAYDKGDIALNEDLSSTWAA
            DTAQITQKWEAAVAGQLRAYLEGTCTVWLRYLENGKETLQRAADPPKXTHVTHP
            ISDHEATLRCWALGFPYPAITLWQPDGDTQTELVTETPAGDRTFKWAAVVPFSG
            EQRVYCHVQHEGLPKPLTLRWPESSQSTPIPIVGIVAGLAVLVV"
BASE COUNT 196 a 310 c 314 g 148 t
ORIGIN

alignment_scores:
  Quality: 50.00      Length: 10
  Ratio: 5.000       Gaps: 0
  Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-23 x HSU11267 ..
Align seg 1/1 to: HSU11267 from: 1 to: 968
1 ArgGluAspLeuArgThrLeuLeuArgTyr 10
|||||
295 CGAGAGGACCTGCGGACCTGCTCGCTAC 324

seq_name: gb_prl:HUMMHC

seq_documentation_block:
LOCUS      HUMMHC 1017 bp mRNA PRI 07-JAN-1995
DEFINITION Human MHC class I HLA-B*27-HS mRNA, 3' end.
ACCESSION  M62852
VERSION     M62852.1 GI:187760
KEYWORDS    cell surface antigen; class I gene; integral membrane protein;
            major histocompatibility complex.
SOURCE      Homo sapiens cDNA to mRNA.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1017)
AUTHORS    Choo,S.Y., Fan,L.A. and Hansen,J.A.
TITLE      A novel HLA-B*27 allele maps B27 allospecificity to the region
            around position 70 in the alpha 1 domain
JOURNAL    J. Immunol. 147 (1), 174-180 (1991)
MEDLINE    91268545
FEATURES    Location/Qualifiers
            source
            1..1017
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /map="6p21.3"
            1..1017
            /gene="HLA-B"
            <1..1017
            /gene="HLA-B"
            /codon_start=1
            /db_xref="GDB:G00-120-048"
            /product="MHC HLA-B*27-HS"
            /protein_id="AAA59647.1"
            /db_xref="GI:187761"
            /translation="GSHSMRYFHTSVSRGGRPEFTVGYVDDTLFVRFSDAASPR

```

EEPRAPWIEQEPYWDRETOICKAKAQTDRDLRLTLRYYNQSEAGSHTLQSMYQCD  
VGPGRLLRGHNOYAYDKDYIALNEDLRSWTAADTAQITORKWEAAAEQLRAYL  
EGECVWLRVYLENGKETLORADPKVTHHPISDHEATLRCHALGFLPAEITLTWO  
RDGQDOTDELTVETRPAGDRTFKWAAVVPVPSGEEQRYTCHVQHGLPKPLTLRWEP  
SSQSTVPVIGIVAGLAVLVVYVIGAAVAAVMCRKSSGGKGGSYQAACSDSAQGSVDV  
SLTA"

BASE COUNT 207 a 308 c 343 g 159 t  
ORIGIN

alignment\_scores:  
Quality: 50.00 Length: 10  
Ratio: 5.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-08-653-294-23 x HUMHMC ..

Align seg 1/1 to: HUMHMC from: 1 to: 1017

1 ArgGluAspLeuArgThrLeuLeuArgTyr 10  
|||||  
223 CGAGAGGACCTGCGGACCTGCTCCGCTAC 252

seq\_name: gb\_pat:A28264

seq\_documentation\_block:  
LOCUS A28264 DNA PAT 24-MAY-1995  
DEFINITION H.sapiens mRNA for HLA-B 27 from patent EP0226069.  
ACCESSION A28264  
VERSION A28264.1 GI:905320  
KEYWORDS  
SOURCE human.

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;  
Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1026)  
AUTHORS Szoets,H., Weiss,E., Doerner,C., Lang,M., Meo,T. and  
Riethmueller,G.

TITLE HLA-B 27, DNA coding therefor and its utilization  
JOURNAL Patent: EP 0226069-A 1 24-JUN-1987;  
Riethmueller, Gert, Prof. Dr

FEATURES  
Location/Qualifiers

Source  
1..1026  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

BASE COUNT 213 a 307 c 344 g 162 t  
ORIGIN

alignment\_scores:  
Quality: 50.00 Length: 10  
Ratio: 5.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-08-653-294-23 x A28264 ..

Align seg 1/1 to: A28264 from: 1 to: 1026

1 ArgGluAspLeuArgThrLeuLeuArgTyr 10  
|||||  
223 CGAGAGGACCTGCGGACCTGCTCCGCTAC 252

seq\_name: gb\_pat:E01342

seq\_documentation\_block:  
LOCUS E01342 RNA PAT 29-SEP-1997  
DEFINITION cDNA encoding C-terminal Fragment of HLA-B27.  
ACCESSION E01342  
VERSION E01342.1 GI:2169599  
KEYWORDS JP 1987228281-A/2.  
SOURCE Homo sapiens.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;  
Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 1026)  
AUTHORS Hannercore,S., Eriabeeto,W., Kurisuta,D., Maagotsuto,R., Tomaso,M.  
and Geruto,R.

TITLE HLA-B27, DNA ENCODING THE SAME AND ITS USE

JOURNAL

Patent: JP 1987228281-A 2 07-OCT-1987;

COMMENT

OS Human

PN JP 1987228281-A/2

PD 07-OCT-1987

PR 28-NOV-1986 JP 1986284078

PF 28-NOV-1985 DE 85 3542024, 21-DEC-1985 DE 85 3545576 PI

PG HANNERCORE SUTSUETSU, ERIZABEETO WAIKU, KURISUTA DERUNAA, PI

MAAGOTSUTO RANGU, TOMASO MEO, GERUTO RIITOMIYURAA PC

C12N15/00,C07H21/04,C12P21/00,C12Q1/68,G01N33/577//A61K39/00, PC

C07K13/00,

PC C07K15/06,(C12P21/00,C12R1:91);

CC strandedness: Double;

CC topology: Linear;

CC hypothetical: No;

CC anti-sense: No;

CC fragment\_type: C-Terminal Fragment;

CC \*source: cell\_type=Leukocyte;

FH Key Location/Qualifiers

FT CDS 1..1026

FT /product="C-terminal fragment(exon 2 to 7)of  
HLA-B27"

FEATURES

Source

1..1026  
/organism="Homo sapiens"

/db\_xref="taxon:9606"

BASE COUNT 213 a 307 c 344 g 162 t

ORIGIN

alignment\_scores:

Quality: 50.00 Length: 10

Ratio: 5.000 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-08-653-294-23 x E01342 ..

Align seg 1/1 to: E01342 from: 1 to: 1026

1 ArgGluAspLeuArgThrLeuLeuArgTyr 10

|||||

223 CGAGAGGACCTGCGGACCTGCTCCGCTAC 252

seq\_name: gb\_pri:HUMMHZUNIA

seq\_documentation\_block:

LOCUS HUMMHZUNIA 1084 bp mRNA PRI 27-SEP-1993

DEFINITION Human MHC class I (HLA-B 27052) mRNA fragment.

ACCESSION L20086

VERSION L20086.1 GI:307282

KEYWORDS class I gene; lymphocyte antigen; major histocompatibility complex.

SOURCE Homo sapiens (strain South American Amerindian) cDNA to mRNA.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;  
Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 1084)

AUTHORS Watkins,D.I., McAdam,S.N., Liu,X., Strang,C.R., Milford,E.L.,  
Levine,C.G., Garber,T.L., Dogon,A.L., Lord,C.I., Ghim,S.H.,  
Troup,G.M., Hughes,A.L. and Letvin,N.L.

TITLE New recombinant HLA-B alleles in a tribe of South American

Amerindians indicate rapid evolution of MHC class I loci

Nature 357, 329-333 (1992)

JOURNAL

MEDLINE

92269956

FEATURES

Location/Qualifiers

```

/db_xref="GI:1144334"
/translation="MRVAPTPTLLLLLWGAVALTWAGSSMRYFHTFSVRPQGGP
RISGVYDQTFVRFSDAASPRTFAPRWIEQEPGWDEQTKINTNYREDL
RTLLYNSDEASHSTLQNNYGCGVGRLLRGHQDAYQKDYIALNEDLSWTA
DTAATQIKRWAAEQRLAYLEGCEVWRVRLYENKGTLRADPTTHPHI
SDHAEITRWALGAFYPAEITLWQDGEQDQITELVTRPAGDTQKAAVVP
EEOQYTCVHGHEGLPKPLTRWEPSSVPIGVIGVLAIVVIGAVVAVMCR
KSSGKGKGSYSOACSDSAOGSDVSLTA"

```



JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1089)  
AUTHORS Seurnyck,K.L. and Baxter-Lowe,L.A.  
TITLE Direct Submission  
JOURNAL Submitted (22-SEP-1997) Molecular Genetics, Richland Memorial  
Hospital, 7 Richland Medical Park, Columbia, SC 29203, USA  
FEATURES  
Source  
1..1089  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="6"  
/cell\_type="EBV transformed B-lymphocytes"  
gene  
1..1089  
/gene="HLA-B"  
/allele="B27052"  
CDS  
1..1089  
/gene="HLA-B"  
/codon\_start=1  
/product="MHC class I antigen"  
/protein\_id="AAC42275.1"  
/db\_xref="GI:3643697"  
/translation="MRVTEPRTLILLGCAVALTETWAGSHSMRYEHTSVSRPGRGEP  
RFITVGYDDTLFVRFSDAASPREEPRAPWIEQGPPEYWDRETOICKAKQTDREDL  
RFLRYINQSEAGSHLQNMVGCVDGPDGRLLRGYHQDAYDKDYIALNEDLUSWTAA  
DTAAQITQRKWEARVAEQRLRAYLEGECEVWLRRLRYLENGKETLQADPPKTHVTHPI  
SDHEATLRCAWLGFPFAEITLTWRDGEDQTDTELVTETRPAGDRTFKRWAAVVVPSG  
EEOYTCRHOEGLEKPLTLRWEPSSQSTVPIVGIVAGLAVLAVVVIGAVVAVMCR  
KSSGGKGSYSQAACSDSAQGSQSVSLTA"

BASE COUNT 217 a 332 C 368 G 172 t  
ORIGIN

alignment\_scores:  
Quality: 50.00 Length: 10  
Ratio: 5.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-08-653-294-23 x AF026218 ..

Align seg 1/1 to: AF026218 from: 1 to: 1089

1 ArgGluAspLeuArgThrLeuLeuArgTyr 10  
|||||  
295 CGAGAGACCTGCGGACCTGCTCCGCTAC 324

**THIS PAGE BLANK (USPTO)**

OM of: US-08-653-294-23 to: N\_Geneseq\_36:\* out\_format: pfs

Date: Feb 8, 2000 7:31 PM

About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:  
-MODEL=frame+2n.model -DEV=xlp  
-O=/cgn1.1/USPTO.spool/US08653294/runat\_04022000.160701\_15807/app\_query.fasta.2  
-DB=N\_Geneseq\_36 -QFMT=fastap -SUFFIX=eng -GAPOP=12.000  
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000  
-GGAPOP=4.500 -GGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500  
-DELOP=6.000 -DELOPEXT=7.000 -XGAPOP=10.000 -XGAPEXT=0.500  
-DELOP=6.000 -DELOPEXT=7.000 -START=1 -MATRIX=blotsum62  
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR\_SCORE=pct  
-ALIGN=15 -MODE=LOCAL -OUTENT=pfs -NORM=ext -MINLEN=0  
-MAXLEN=1000000 -USER=US08653294 -NCPU=6 -ICPU=3 -NO\_XLPXY -WAIT  
-THREADS=1

Search information block:

Query: US-08-653-294-23

Query length: 10

Database: N\_Geneseq\_36:\*

Database sequences: 311585

Database length: 125096042

Search time (sec): 873.190000

score\_list:

Sequence	Strd Orig	zscore	EScore	len	Documentation
N_Geneseq_36:N70935	+ 50.00	169.98	0.0723	1026	Sequence encoding the human h
N_Geneseq_36:N70225	+ 50.00	158.36	0.3207	3874	Sequence of genomic DNA encodi
N_Geneseq_36:T61639	+ 50.00	153.76	0.5784	6553	HLA B27 consensus sequence. De
N_Geneseq_36:V35645	+ 37.00	113.95	95.46	2386	Cladosporium oxysporum glucose
N_Geneseq_36:V21209_16	- 37.00	85.05	3.9e+03	64976	Continuation (17 of 17) of
N_Geneseq_36:V89109	- 36.00	124.64	24.22	458	EST clone B2187. New polynucle
N_Geneseq_36:V17143	- 36.00	119.20	48.66	853	Human proteolipid (PLHu) cDNA.
N_Geneseq_36:V13170	- 36.00	97.20	81.90	10555	Enterococcus faecalis genome
N_Geneseq_36:V28651	+ 35.00	116.10	72.45	793	Ripening banana pulp cDNA clone
N_Geneseq_36:Q29167	+ 34.00	121.78	34.96	270	HLA-Bw 52 exon 2 alpha-1 domain
N_Geneseq_36:Q01834	+ 34.00	109.61	166.61	1086	Sequence encoding HLA-B*51 anti
N_Geneseq_36:Q01822	+ 34.00	109.61	166.61	1086	Sequence encoding HLA-B*52 ant
N_Geneseq_36:Q05693	+ 34.00	109.58	167.13	1089	HLA-B*51 gene for production of
N_Geneseq_36:Q05701	+ 34.00	109.58	167.13	1089	HLA-B*52 gene for production of
N_Geneseq_36:Q01214	+ 34.00	109.58	167.13	1089	HLA-B*53 exon. HLA-B*53 gene,
N_Geneseq_36:X04639	+ 34.00	100.76	517.95	2985	Nucleic acid encoding an enzym
N_Geneseq_36:V22816	- 34.00	96.16	934.82	5053	Nucleotide sequence of the spe
N_Geneseq_36:V17236	- 34.00	96.16	935.24	5055	DNA from a region of NP strain
N_Geneseq_36:T00770	+ 34.00	92.53	1.5e+03	7653	Sirolimus effector protein gen
N_Geneseq_36:T33872	+ 34.00	92.34	1.5e+03	7824	Human RAPT1 cDNA. Rapamycin bi
N_Geneseq_36:V28518	+ 34.00	92.20	1.6e+03	7943	FRAP (tor1) cDNA. Selective in
N_Geneseq_36:V21209_00	+ 34.00	69.22	2.9e+04	110000	Methanococcus jannaschii c
N_Geneseq_36:X20248_07	- 34.00	69.22	2.9e+04	110000	Continuation (8 of 10) of
N_Geneseq_36:X13461	+ 33.00	99.12	639.73	2349	Enterococcus faecalis genome c
N_Geneseq_36:V60573	+ 33.00	95.83	975.01	3420	Cercospora nicotianae cercospo
N_Geneseq_36:V52088	+ 32.50	98.71	673.66	1986	Helicobacter polypeptide GHPO
N_Geneseq_36:X14516	+ 32.50	98.71	673.66	1986	H. pylori GHPO 1278 gene. New
N_Geneseq_36:T68022	+ 32.50	98.69	673.94	1992	H. pylori inner membrane prote
N_Geneseq_36:X30428	+ 32.50	98.53	689.66	2038	H. pylori outer membrane prote
N_Geneseq_36:T22846	+ 32.00	125.98	20.41	71	Human gene signature HUMGS04520
N_Geneseq_36:T96627	+ 32.00	108.94	181.45	498	cDNA encoding B. napus cytosoli
N_Geneseq_36:T91186	+ 32.00	108.27	197.88	538	Flea saliva protein fspN(O) cDN
N_Geneseq_36:V73403	+ 32.00	108.27	197.88	538	Flea saliva protein nucleic aci
N_Geneseq_36:T79328	- 32.00	102.70	404.21	1017	DNA encoding Terebinthabacter es
N_Geneseq_36:T03250	- 32.00	97.73	764.07	1794	Enterobacter sp. S262 sucrose
N_Geneseq_36:V22706	- 32.00	96.16	934.63	2147	Mouse recombinant hsrEC2 gene
N_Geneseq_36:Q37687	- 32.00	94.69	1.1e+03	2540	Rat choline kinase gene. Rat c
N_Geneseq_36:T36567	- 32.00	94.05	1.2e+03	2734	Thermococcus barosii DNA poly
N_Geneseq_36:V10362	+ 32.00	90.17	2.0e+03	4257	Infected cell protein number 4
N_Geneseq_36:V68520	+ 32.00	90.17	2.0e+03	4257	The nucleotide sequence of the
N_Geneseq_36:N91312	+ 32.00	85.22	3.8e+03	7499	pCHL1 plasmid. Synthetic oligo

N\_Geneseq\_36:Q27429 + 32.00 85.22 3.8e+03 7502 ! pCTD. PCTD plasmid from Chl  
N\_Geneseq\_36:T85392 + 32.00 83.72 4.6e+03 8906 ! Human TRIO phosphoprotein c  
N\_Geneseq\_36:Q76213 - 32.00 81.11 6.4e+03 12001 ! HSV L/ST region. Herpes si  
N\_Geneseq\_36:V31206 + 32.00 80.85 6.6e+03 12368 ! E. coli J96 pathogenicity

seq\_name: N\_Geneseq\_36:N70935

seq\_documentation\_block:

ID N70935 standard; DNA; 1026 BP.

AC N70935;

DT 10-APR-1991 (first entry)

DE Sequence encoding the human histocompatibility antigen HLA B27.

KW Rheumatic disorder; genetic screening; diagnosis;

KW Ankylosing spondylitis; ss.

OS Homo sapiens.

FH Key Location/Qualifiers

FT Cds 1..1026

PD DE3542024-A.

PF 04-JUN-1987.

PF 28-NOV-1985; 542024.

PR 28-NOV-1985; DE-542024.

PR 21-DEC-1985; DE-545576.

PA (BEHW ) BEHRINGERWERKE AG.

PI Riethmuller G, Meo T, Weiss E, Szots H;

DR WPI; 87-157893/23.

DR P-PSDB; P70590.

PT DNA coding for antigen HLA B27 - and diagnostic reagents contg.

PT such DNA, antigen or antibody

PS Claim 2; Page 4; 5pp; German.

CC The DNA may be used as a hybridisation probe for detecting the HLA

CC B27 gene, e.g. for assessing susceptibility to rheumatic disorders

CC such as ankylosis spondylitis, or may be used to transform cells

CC for prodn. of HLA B27. The HLA B27 may be used to detect HLA B27

CC antibody in human serum, or to produce mono- or polyclonal HLA B27

CC antibodies for use in immunoassay.

CC Sequence 1026 BP; 213 A; 307 C; 344 G; 162 T;

alignment\_scores:

Quality: 50.00 Length: 10

Ratio: 5.000 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-08-653-294-23 x N70935 ..

Align seg 1/1 to: N70935 from: 1 to: 1026

1 ArgGluAspLeuArgThrLeuLeuArgTyr 10

|||||

223 CGAGAGGACCTGGGACCCCTGCTCCGCTAC 252

seq\_name: N\_Geneseq\_36:N70225

seq\_documentation\_block:

ID N70225 standard; DNA; 3874 BP.

AC N70225;

DT 03-APR-1991 (first entry)

DE Sequence of genomic DNA encoding human histocompatibility antigen

DE HLA-B 27.

KW Ankylosing spondylitis; rheumatic disorder; diagnosis; ss.

OS Homo sapiens.

FH Key Location/Qualifiers

FT Intron 518..590

/\*tag= a

720..989

/\*tag= b

1090..1506

/\*tag= c

1932..2357

/\*tag= d

2450..2566

/\*tag= e

```

FT intron 3009..3041
FT /*tag= f
FT intron 3148..3191
FT /*tag= g
PN EP-226069-A.
PD 24-JUN-1987.
PF 21-NOV-1986; 116139.
PR 01-JAN-1985; DE-542024.
PR 21-DEC-1985; DE-545576.
PA (BEHW ) BEHRINGER AG.
PI Szöts H, Weiss E, Dörner C, Lang M, Meo T, Riethmüller G;
DR WPI: 87-171469/25.
DR P-PSDB; P70155.
DR DNA coding for human histocompatibility antigen HLA-B 27 - useful
PT for diagnosis and antigen and antibody prodn.
PS Claim 1; p6; 13pp; German.
CC The DNA may be used to detect the HLA-B 27 gene (opt. mutated) in
CC human genetic material. The HLA-B 27 may be used to detect anti-HLA-
CC B 27 antibodies in human serum. The antibodies may be used to
CC determine HLA-B 27 levels in human serum, e.g. for diagnosis of
CC rheumatic disorders, esp. ankylosing spondylitis.
SQ Sequence 3874 BP; 751 A; 1094 C; 1171 G; 858 T;

alignment_scores:
  Quality: 50.00 Length: 10
  Ratio: 5.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-23 x N70225 ..
Align seg 1/1 to: N70225 from: 1 to: 3874

1 ArgGluAspLeuArgThrLeuLeuArgTyr 10
|||||
941 CGAGAGGACCTCGGACCTGCTCGCTAC 970

seq_name: N_Geneseq_36:T61639

seq_documentation_block:
ID T61639 standard; DNA; 6553 BP.
AC T61639;
DT 05-JUN-1997 (first entry)
DE HLA B27 consensus sequence.
KW HLA B27; seronegative spondylarthropathy; ankylosing spondylitis;
KW Reiter's syndrome; arthritis; acute anterior uveitis; diagnosis;
KW ss; ds.
OS Homo sapiens.
FH Key Location/Qualifiers
FT mrna 3968..6653
FT /*tag= a
FT /note= "HLA-B27 3' flanking region, downstream of
FT 3' untranslated region"
FT mrna 4112..4556
FT /*tag= b
FT /note= "3' flanking region diagnostic for genetic
FT predisposition to SNSA"
FT mrna 4270..4556
FT /*tag= b
FT /note= "3' flanking region diagnostic for genetic
FT predisposition to SNSA"
FT misc_difference 4495
FT /*tag= d
FT /note= "absence of cytosine at this site is
FT indicative of a predisposition to SNSA"
PN WO9709450-A1.
PD 13-MAR-1997.
PF 16-AUG-1996; U13256.
PR 01-SEP-1995; US-522942.
PA (CEDA-) CEDARS SINAI MEDICAL CENT.
PI Tyan DB;
DR WPI: 97-192924/17.

```

```

PT Detecting pre-disposition to seronegative spondylarthropathies -
PT from the absence of a C residue at a specific position in the
PT 3'-flanking region of the HLA B27 allele
PS Claim 1; Page 52-56; 68pp; English.
CC Genetic predisposition to seronegative spondylarthropathies (SNSA)
CC is detected by determining the absence of a cytosine nucleotide in
CC the 3' flanking region (see also T61647-48) of an HLA-B gene at a
CC position corresponding to nucleotide 4495 of the HLA-B27 consensus
CC sequence given in T61639. Probes and primers (see also T61640-46)
CC based on this region can be used in diagnostic assays to detect the
CC genetic predisposition to SNSA, and permit the distinction of B27+
CC individuals who are resistant to SNSA from B27- normal individuals
CC who are susceptible (but as yet unaffected) to such diseases.
SQ Sequence 6553 BP; 1443 A; 1619 C; 2017 G; 1474 T;

alignment_scores:
  Quality: 50.00 Length: 10
  Ratio: 5.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-23 x T61639 ..
Align seg 1/1 to: T61639 from: 1 to: 6553

1 ArgGluAspLeuArgThrLeuLeuArgTyr 10
|||||
1102 CGAGAGGACCTCGGACCTGCTCGCTAC 1131

seq_name: N_Geneseq_36:V35645

seq_documentation_block:
ID V35645 standard; DNA; 2386 BP.
AC V35645;
DT 08-SEP-1998 (first entry)
DE Cladosporium oxysporum glucose oxidase encoding DNA.
KW Glucose oxidase; Cladosporium oxysporum; enzyme; bread improver;
KW antimicrobial agent; toothpaste; detergent; stain removal;
KW dough additive; hydrogen peroxide generator; ss.
OS Cladosporium oxysporum.
FH Key Location/Qualifiers
FT 5'UTR 1..350
FT /*tag= a
FT CDS 351..2192
FT /*tag= b
FT /product= "glucose oxidase"
FT sig_peptide 351..419
FT /*tag= c
FT /note= "Predicted signal sequence"
FT mat_peptide 420..2189
FT /*tag= d
FT 3'UTR 2190..2386
FT /*tag= e
PN WO9820136-A1.
PD 14-MAY-1998.
PF 03-NOV-1997; U20174.
PR 07-NOV-1996; US-746257.
PA (NOVO ) NOVO NORDISK BIOTECH INC.
PA (NOVO ) NOVO NORDISK AS.
PI Berka RM, Cherry JR, Halkier T;
DR WPI: 98-286952/25.
DR P-PSDB; W60593.
DR New nucleic acid encoding glucose oxidase active at acidic pH, from
PT Cladosporidium - and related vectors and host cells, producing
PT enzyme useful as bread improver, antimicrobial additive for
PT toothpaste, etc. and hydrogen peroxide generator in detergents
PS Claim 4; Fig 2A-B; 83pp; English.
CC This DNA encodes a Cladosporium oxysporum glucose oxidase. Host cells
CC containing a construct comprising the glucose oxidase encoding nucleic
CC acid sequence with regulatory sequences are used to produce recombinant
CC glucose oxidase. Preferred fragments of the nucleic acid are present in
CC pGEX4A and pGEX6A, contained in E. coli deposited as NRRL B-21628 and

```

CC B-21629. The host cells are particularly Fusarium. The recombinant  
 CC glucose oxidase is useful as dough additive to improve the gluten  
 CC quality. It can be used as additive for toothpaste (particularly used  
 CC with a thiocyanate and lactoperoxidase to generate antimicrobial  
 CC oxythiocyanate anion), mouthwash, denture cleaners, soaps, hair and body  
 CC care products. It can be added to cleaning solutions for contact lenses  
 CC as bleach or as an antibacterial agent. It is also useful as a hydrogen  
 CC peroxide generator in laundry and dishwashing detergents, particularly  
 CC for stain removal.  
 SQ Sequence 2386 BP; 583 A; 707 C; 581 G; 515 T;

alignment\_scores:  
 Quality: 37.00 Length: 9  
 Ratio: 4.625 Gaps: 0  
 Percent Similarity: 88.889 Percent Identity: 88.889

alignment\_block:

US-08-653-294-23 x V35645 ..

Align seg 1/1 to: V35645 from: 1 to: 2386

1 ArgGluAspLeuArgThrLeuLeuArg 9  
 |||||  
 605 CGTGAGGACTTACGGACAGCCTCGA 631

seq\_name: N\_Geneseq\_36:V21209\_16

seq\_documentation\_block:

Continuation (17 of 17) of V21209 from base 1600001 (Methanococcus jannaschii circular  
 WP Sequence split into 17 fragments LOCUS V21209 Accession V21209

Fragment Name	Begin	End
WP V21209_00	1	110000
WP V21209_01	100001	210000
WP V21209_02	200001	310000
WP V21209_03	300001	410000
WP V21209_04	400001	510000
WP V21209_05	500001	610000
WP V21209_06	600001	710000
WP V21209_07	700001	810000
WP V21209_08	800001	910000
WP V21209_09	900001	1010000
WP V21209_10	1000001	1110000
WP V21209_11	1100001	1210000
WP V21209_12	1200001	1310000
WP V21209_13	1300001	1410000
WP V21209_14	1400001	1510000
WP V21209_15	1500001	1610000
WP V21209_16	1600001	1664976

alignment\_scores:  
 Quality: 37.00 Length: 10  
 Ratio: 4.111 Gaps: 0  
 Percent Similarity: 90.000 Percent Identity: 70.000

alignment\_block:

US-08-653-294-23 x V21209\_16/rev ..

Align seg 1/1 to: reverse of: V21209\_16 from: 1 to: 64976

1 ArgGluAspLeuArgThrLeuLeuArgTyr 10  
 |||||  
 49695 AGAGAAAATAATTAAGACAGTGCTAAGATAT 49666

seq\_name: N\_Geneseq\_36:V89109

seq\_documentation\_block:

ID V89109 standard; cDNA: 458 BP.

AC V89109;

DT 15-FEB-1999 (first entry)

DE: EST clone B2187.

KW Human; secreted protein; expressed sequence tag; EST; haematopoiesis;

KW tissue growth; activin; inhibin; chemotaxis; chemokinesis; haemostatic;  
 KW receptor; ligand; thrombolytic; anti-inflammatory; cadherin; anti-tumour;  
 KW gene therapy; ss.  
 OS Homo sapiens.  
 PN WO9845436-A2.  
 PD 15-OCT-1998.  
 PF 10-APR-1998; U06955.  
 PR 10-APR-1997; US-838821.  
 PA (GENY ) GENETICS INST INC.  
 PI Agostino M, Jacobs K, Lavallie ER, McCoy JM, Merberg D,  
 PI Racine LA, Spaulding V, Treacy M;  
 DR WPI; 99-070077/06.  
 PT New polynucleotides encoding human secreted proteins - derived from  
 PT e.g. human blood, kidney, foetal lung, placenta, testes, brain,  
 PT ovary, pituitary, retina and colon cDNA libraries.  
 PS Claim 1; Page 113; 618pp; English.  
 CC The present sequence represents a human expressed sequence tag (EST).  
 CC The polynucleotide, which is a secreted EST, and the encoded protein  
 CC are predicted to have useful biological activities which would make  
 CC them suitable for treating, preventing or ameliorating medical  
 CC conditions in humans and animals, although no supporting data is  
 CC given. Suggested activities include nutritional activity, immune  
 CC stimulating or suppressing activity, haematopoiesis regulating  
 CC activity, tissue growth activity, activin/inhibin activity,  
 CC chemotactic/chemokinetic activity, haemostatic and thrombolytic  
 CC activity, receptor/ligand activity, anti-inflammatory activity,  
 CC cadherin/tumour invasion suppressor activity, tumour inhibition  
 CC activity. The polynucleotide may also be useful for gene therapy.  
 SQ Sequence 458 BP; 132 A; 70 C; 74 G; 162 T;

alignment\_scores:  
 Quality: 36.00 Length: 9  
 Ratio: 4.000 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 66.667

alignment\_block:

US-08-653-294-23 x V89109/rev ..

Align seg 1/1 to: reverse of: V89109 from: 1 to: 458

2 GluAspLeuArgThrLeuLeuArgTyr 10  
 |||||  
 358 GAAGACATTAAATACATTAAAGATAC 332

seq\_name: N\_Geneseq\_36:V17143

seq\_documentation\_block:

ID V17143 standard; cDNA: 853 BP.

AC V17143;

DT 07-JUL-1998 (first entry)

DE Human proteolipid (PLHu) cDNA.

KW Human proteolipid; PLHu; exocytosis; membrane trafficking;

KW rat plasmolipin; chemokine; inflammation; ss.

OS Homo sapiens.

FH Key

FT Location/Qualifiers

FT CDS

FT 19..480

FT /\*tag= a

FT /product= "Human proteolipid"

PN WO9804691-A1.

PD 05-FEB-1998.

PF 21-JUL-1997; U12734.

PR 26-JUL-1996; US-695736.

PA (INCY-) INCYTE PHARM INC.

PI Au-Young J, Bandman O, Goli SK, Hillman JL;

DR WPI; 98-130683/12.

DR P-PSDB; W5217.

PT Human proteolipid PLHu - used to accelerate diagnosis and proper

PT treatment of diseases and conditions associated with abnormal

PT membrane trafficking

PS Claim 4; Fig 1A-1B; 53pp; English.

CC The present sequence is of a human proteolipid (PLHu) cDNA which

CC was partially isolated as Incyte clone 640699 derived from a human breast

CC cDNA library. The invention provides PLHu cDNA for expression of PLHu  
 CC protein. Homology of PLHu protein to rat plasmalipin indicates its  
 CC involvement in exocytosis. Exocytosis facilitated by PLHu may influence  
 CC membrane trafficking within the cell and could affect the release of  
 CC chemokines involved in cell migration, proteases which are active in  
 CC inflammation or other similar activities involving endothelial cells,  
 CC fibroblasts, etc. The invention also claims for antibodies against  
 CC PLHu which can be used for e.g. in diagnostic tests to accelerate  
 CC diagnosis and proper treatment of conditions associated with abnormal  
 CC membrane trafficking.  
 SQ Sequence 853 BP; 233 A; 187 C; 162 G; 271 T;

alignment\_scores:  
 Quality: 36.00 Length: 9  
 Ratio: 4.000 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 66.667

alignment\_block:

US-08-653-294-23 x V17143/rev ..

Align seg 1/1 to reverse of: V17143 from: 1 to: 853

2 GluAspLeuArgThrLeuLeuArgTyr 10  
 |||||:|||||:|||||:|||||  
 709 GAAGACATTATACATTATTAAGATAC 683

seq\_name: N\_Geneseq\_36:X13170

seq\_documentation\_block:

ID X13170 standard; DNA; 10555 BP.

AC X13170;

DT 19-MAR-1999 (first entry)

DE Enterococcus faecalis genome contig SEQ ID NO:233.

KW Enterococcus faecalis; contig; detection; Enterococcal infection;

OS vaccine; attenuation; computer readable medium; ds.

PN Enterococcus faecalis.

PN WO9850555-A2.

PD 12-NOV-1998.

PF 04-MAY-1998; U08985.

PR 14-NOV-1997; US-066009.

PR 06-MAY-1997; US-044031.

PR 16-MAY-1997; US-046655.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Barash SC, Dillon PJ, Kunsch CA;

DR WPI; 99-045171/04.

PT New isolated Enterococcus faecalis polynucleotides and polypeptides  
 PT - used to develop products for the detection of Enterococcus and for  
 PT use in vaccines for prevention or attenuation of Enterococcus  
 PT infection.

PS Claim 1; Page 1164-1169; 2084pp; English.

CC A computer readable medium has been developed which has recorded on it  
 CC 982 nucleotide sequences isolated from the Enterococcus faecalis genome.  
 CC X12938 to X13919 represent these nucleotide sequences which are primary  
 CC nucleotide sequences, also known as contigs. The computer-based system  
 CC can identify fragments of the Enterococcus faecalis genome with  
 CC commercial importance. The products can be used to detect the presence  
 CC of Enterococcus faecalis in samples. They can also be used for  
 CC diagnosing Enterococcal infection in an animal and monitoring  
 CC progression of disease, and for identifying agents which can be used to  
 CC modulate the growth or pathogenicity of Enterococcus faecalis, or  
 CC another related organism, in vivo or in vitro. In particular the  
 CC polypeptides encoded by the Enterococcus faecalis nucleotide sequences  
 CC can be used in vaccines to prevent or attenuate an Enterococcal  
 CC infection.

SQ Sequence 10555 BP; 2903 A; 2453 C; 1793 G; 3396 T;

alignment\_scores:  
 Quality: 36.00 Length: 9  
 Ratio: 4.500 Gaps: 0  
 Percent Similarity: 88.889 Percent Identity: 77.778

alignment\_block:

US-08-653-294-23 x X13170 ..

Align seg 1/1 to: X13170 from: 1 to: 10555

2 GluAspLeuArgThrLeuLeuArgTyr 10

:::|||||:|||||:|||||:|||||

5890 CAAGATTTAAGGAAGTTGTTGCGTAT 5916

seq\_name: N\_Geneseq\_36:V28651

seq\_documentation\_block:

ID V28651 standard; cDNA; 793 BP.

AC V28651;

DT 29-JUL-1998 (first entry)

DE Ripening banana pulp cDNA clone U-U131 SEQ ID NO:9.

KW Banana; ripening; pulp; Musa acuminata cv. Grand Nain; fruit;

KW genetic control; tissue senescence; ss.

OS Musa acuminata.

PN WO9811228-A2.

PD 19-MAR-1998.

PF 08-SEP-1997; G02424.

PR 25-APR-1997; GB-008366.

PR 10-SEP-1996; GB-018862.

PA (ZENE ) ZENECA LTD.

PI Bird CR, Medina-Suarez RDJ, Seymour GB;

DR WPI; 98-207389/18

PT Modulation of ripening or tissue senescence in bananas - comprises  
 PT use of DNA isolated from ripening banana pulp to produce genetically  
 PT modified fruit

PS Claim 1; Page 23-24; 72pp; English.

CC The present sequence represents a cDNA clone isolated from ripening  
 CC banana pulp. 57 clones were isolated and are given in V28643 to V28699.  
 CC The cDNA clone sequences can be used in a method of modulating ripening  
 CC or tissue senescence process in plants of the genus Musa. The method  
 CC comprises: (a) inserting into the plant material at least 1 of the 57  
 CC sequences (as above); (b) regenerating the plant material, and (c)  
 CC selecting from the transformed regenerants, plants with modulated  
 CC ripening or tissue senescence characteristics. Also described in the  
 CC present invention are: (1) plants, their progeny, seed and material  
 CC obtained from the plants, produced by the above method; (2) a vector  
 CC functional in plants comprising a promoter region which is operably in  
 CC plant cells, a polynucleotide sequence as defined above, and a  
 CC transcription termination sequence; and (3) a method of controlling  
 CC plant pathogens comprising the application of anti-pathogenic agent to  
 CC plants of (1).  
 SQ Sequence 793 BP; 161 A; 275 C; 199 G; 137 T;

alignment\_scores:

Quality: 35.00 Length: 10  
 Ratio: 4.375 Gaps: 0  
 Percent Similarity: 80.000 Percent Identity: 70.000

alignment\_block:

US-08-653-294-23 x V28651 ..

Align seg 1/1 to: V28651 from: 1 to: 793

1 ArgGluAspLeuArgThrLeuLeuArgTyr 10

||| ||| |||||:|||||:|||||

626 CGCCAGATNAAGAACCTCTCGCTTT 655

seq\_name: N\_Geneseq\_36:Q29167

seq\_documentation\_block:

ID Q29167 standard; DNA; 270 BP.

AC Q29167;

DT 09-MAR-1993 (first entry)

DE HLA-B\*52 exon 2 alpha-1 domain.

KW Human leukocyte antigen; transgenic; germ cells; somatic cells;

KW expression; ss.

PN J04091731-A.

```

PD 25-MAR-1992.
PF 03-AUG-1990; 207329.
PA (OLYU ) OLYMPUS OPTICAL CO.
DR WPI; 92-342893/42.
PT Transgenic non-human mammalian HLA-Bw 52 gene - useful for
PT analysis of expression of gene structure, and prodn. of
PT mouse model of human disease
PS Disclosure: Fig 1: 8pp; Japanese.
CC The sequence shows the exon 2 alpha-1 domain of the human leukocyte
CC antigen-Bw 52 gene. The complete gene may be introduced into non-
CC human mammals, pref. rat or mouse, or their ancestors at the primary
CC developmental biological step via transplantation into the zygote or
CC embryo to generate transgenic non-human mammals incorporating the
CC HLA-Bw 52 gene in both their germ cells and somatic cells. Transgenic
CC non-human mammals contg. HLA-Bw 52 are useful for the analysis of
CC expression of the gene, its structure, and prodn. of mouse models of
CC human disease. See also Q29166-72.
SQ Sequence 270 BP; 59 A; 88 C; 86 G; 37 T;

alignment_scores:
  Quality: 34.00 Length: 10
  Ratio: 4.250 Gaps: 0
  Percent Similarity: 80.000 Percent Identity: 70.000
alignment_block:
US-08-653-294-23 x Q29167 ..
Align seg 1/1 to: Q29167 from: 1 to: 270

1 ArgGluAspLeuArgThrLeuLeuArgTyr 10
|||||:||||| |||||
222 CGAGAGAACCTGCGGATCGCGCTCGCTAC 251

seq_name: N_Geneseq_36:Q01834
seq_documentation_block:
ID Q01834 standard; DNA; 1086 BP.
AC Q01834;
DT 19-MAR-1991 (first entry)
DE Sequence encoding HLA-B51 antigen.
KW Probe: HLA class I DNA; immunogen; ss.
OS Homo sapiens.
PN EP354580-A.
PD 14-FEB-1990.
PF 10-AUG-1989.
PR 11-AUG-1988; JP-200758.
PA (OLYU) Olympus Optical Co., Ltd.
PI Kano K, Takiguchi;
DR WPI; 90-046289/07.
PT New DNA for class 1 human leukocyte antigens and derived probes and
PT transformed cells, useful for DNA typing, as immunogens etc.
PS Claim 1; Page 11; 23pp; English.
CC The HLA class I DNA can be used a source of probes for use in DNA
CC typing. Transformed cells, which are useful as immunogens, can be
CC obtained by introducing these DNAs into eucaryotic cells.
SQ Sequence 1086 BP; 224 A; 334 C; 356 G; 172 T;

alignment_scores:
  Quality: 34.00 Length: 10
  Ratio: 4.250 Gaps: 0
  Percent Similarity: 80.000 Percent Identity: 70.000
alignment_block:
US-08-653-294-23 x Q01834 ..
Align seg 1/1 to: Q01834 from: 1 to: 1086

1 ArgGluAspLeuArgThrLeuLeuArgTyr 10
|||||:||||| |||||
294 CGAGAGAACCTGCGGATCGCGCTCGCTAC 323

seq_name: N_Geneseq_36:Q01822
seq_documentation_block:
ID Q01822 standard; DNA; 1086 BP.
AC Q01822;
DT 19-MAY-1991 (first entry)
DE Sequence encoding HLA-Bw52 antigen.
KW Probe: HLA class I DNA; immunogen; ss.
OS Homo sapiens.
PN EP-354580-A.
PD 14-FEB-1990.
PF 10-AUG-1989.
PR 11-AUG-1988; JP-200758.
PA (OLYU) Olympus Optical Co., Ltd.
PI Kano K, Takiguchi;
DR WPI; 90-046289/07.
PT New DNA for class 1 human leukocyte antigens and derived probes and
PT transformed cells, useful for DNA typing, as immunogens etc.
PS Claim 2; pp11-12; 23pp; English.
CC The HLA class I DNA can be used a source of probes for use in DNA
CC typing. Transformed cells, which are useful as immunogens, can be
CC obtained by introducing these DNAs into eucaryotic cells.
SQ Sequence 1086 BP; 223 A; 335 C; 358 G; 170 T;

alignment_scores:
  Quality: 34.00 Length: 10
  Ratio: 4.250 Gaps: 0
  Percent Similarity: 80.000 Percent Identity: 70.000
alignment_block:
US-08-653-294-23 x Q01822 ..
Align seg 1/1 to: Q01822 from: 1 to: 1086

1 ArgGluAspLeuArgThrLeuLeuArgTyr 10
|||||:||||| |||||
294 CGAGAGAACCTGCGGATCGCGCTCGCTAC 323

seq_name: N_Geneseq_36:Q05693
seq_documentation_block:
ID Q05693 standard; DNA; 1089 BP.
AC Q05693;
DT 03-JAN-1991 (first entry)
DE HLA-B51 gene for production of monoclonal antibodies.
KW Allootype specific monoclonal anti-HLA antibodies; hybridomas;
KW transgenic animals; HLA-B51 gene; ss.
FH Key Location/Qualifiers
FT exon 1..73
FT FT /tag= a
FT FT /number=1
FT FT /tag= b
FT FT /number=2
FT FT /note="alpha 1-domain"
FT FT 344..619
FT FT /tag= c
FT FT /number=3
FT FT /note="alpha 2-domain"
FT FT 620..895
FT FT /tag= d
FT FT /number=4
FT FT /note="alpha 3-domain"
FT FT 896..1012
FT FT /tag= e
FT FT /number=5
FT FT 1013..1042

```

```

FT      exon      1043..1089
FT      FT      /*tag= g
FT      FT      /number=7
FT      PN      EP-383183-A.
FT      PD      22-AUG-1990.
FT      PF      07-FEB-1990; 102424.
FT      PR      08-FEB-1989; JP-029313.
FT      PA      (OLYU ) OLYMPUS OPTICAL KK.
FT      PI      Takiguchi M;
FT      DR      WPI; 90-255479/34.
FT      PT      Allotype specific monoclonal anti- HLA antibodies prodn. - using
FT      PT      hybridomas derived from transgenic animals carrying HLA gene and
FT      PT      immunised with HLA antigen of different allotype
FT      PS      Disclosure; Fig 1 A-G; 20pp; English.
FT      CC      The human HLA-Bw52 gene was introduced into mouse L cells and
FT      CC      then these cells used to immunise one of the transgenic mice
FT      CC      (See Q05693).
FT      CC      The spleen lymphocytes were fused with myeloma cells (P3x63-Ag8.653).
FT      CC      Hybridomas producing antibodies were selected.
FT      CC      See also Q05701.
FT      SQ      Sequence 1089 BP; 224 A; 335 C; 357 G; 173 T;

alignment_scores:
Quality: 34.00 Length: 10
Ratio: 4.250 Gaps: 0
Percent Similarity: 80.000 Percent Identity: 70.000

alignment_block:
US-08-653-294-23 x Q05693 ..
Align seg 1/1 to: Q05693 from: 1 to: 1089
1 ArgGluAspLeuArgThrLeuLeuArgTyr 10
|||||:|||||
295 CGAGAGAACTGCGGATCGGCTCCGCTAC 324

seq_name: N_Geneseq_36:Q05701

seq_documentation_block:
ID Q05701 standard; DNA; 1089 BP.
AC Q12114;
AT Q12114;
DT 29-AUG-1991 (first entry)
DE HLA-Bw53 exon.
KW Human leukocyte antigen; probe; major histocompatibility complex;
KW MHC; class I; ss.
OS Homo sapiens.
FH Key
FT Location/Qualifiers
FT cds
FT J03112487-A.
FT 14-MAY-1991.
FT 22-SEP-1989; 247697.
FT 22-SEP-1989; JP-247697.
FT PA (OLYU ) OLYMPUS OPTICAL KK.
FT DR WPI; 91-182991/25.
FT DR P-PSDB; R12463.
FT PT HLA-Bw53 gene. DNA probe and transformant cells - used for
FT PT immunisation, identifying specificity of antiserum etc.
FT PS Claim 1; Page 1; lipp; Japanese.
FT CC Probes comprising part of the sequence can be used to identify
FT CC Class I genes. The DNA can be expressed for immunisation of
FT CC animals and prodn. of monoclonal antibodies specific for the
FT CC HLA-Bw53 antigen. See also J03112485 and J03112486.
FT CC
FT SQ Sequence 1089 BP; 222 A; 337 C; 356 G; 174 T;

alignment_scores:
Quality: 34.00 Length: 10
Ratio: 4.250 Gaps: 0
Percent Similarity: 80.000 Percent Identity: 70.000

alignment_block:
US-08-653-294-23 x Q12114 ..

```



Align seg 1/1 to: Q12114 from: 1 to: 1089

1 ArgGluAspLeuArgThrLeuLeuArgTyr 10

|||||:|||||

|||||

295 CGAGAGAACCTGCGGATCGCGCTCCGCTAC 324

**THIS PAGE BLANK (USPTO)**

OM of: US-08-653-294-23 to: EST:\* out\_format : pfs

Date: Feb 8, 2000 6:22 AM

About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEM=framet.p2n.model -DEV=xl  
-O/cgpl1/USPTO.spool/US08653294/runat\_04022000\_160700\_15770/app\_query.fasta.2  
-DB=EST -QMT=fastap -SUFFIX=rst -GAPOP=12.000 -GAPEXT=4.000  
-MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -GAPOP=4.500  
-GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000  
-FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000  
-DELEXT=7.000 -START=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pfs -NORM=ext -MINLEN=0 -MAXLEN=100000 -USER=US08653294  
-NCPU=6 -ICPU=3 -NO\_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-08-653-294-23

Query length: 10

Database: EST:\*

Database sequences: 4538634

Database length: 1887831982

Search time (sec): 7600.090000

score\_list:

Sequence	Strd	Orig	zScore	Escore	Len	Documentation
gb_est38:AL036690	+	50.00	190.64	0.0747	171	! AL036690 DKF2P564D2463_r1 564
gb_gss8:AO039738	+	44.00	160.21	3.70	380	! AO039738 CIT-HSP-2317E17.TF CIT
gb_gss13:AA040598	-	44.00	159.11	4.26	430	! AQ440598 HS_5089_B1_c10.SP6E RP
gb_est20:AA879637	+	43.00	157.12	5.50	347	! AA879637 vs3805.r1 Stratagene
gb_est16:AA596937	+	43.00	154.70	7.51	456	! AA596937 vs4104.r1 Barstead m
gb_est44:AA592217	+	43.00	150.97	12.11	468	! AA592217 vs15e03.r1 Barstead m
gb_est46:AA556868	+	43.00	148.40	16.83	249	! AA556868 UI-R-C2p-ro-d-11-0-UI
gb_est23:AA124815	+	40.00	143.89	30.04	414	! AA124815 am5606.x1 Johnston fi
gb_gss5:AAQ79881	-	40.00	142.63	35.30	477	! AA79881 HS_2338_B1_A11.MR CIT
gb_est25:AA1268994	+	40.00	142.14	37.58	504	! AA126894 g367r01.x1 NCI.CGAP.R
gb_gss15:AA060634	+	40.00	141.25	42.12	557	! AA060634 HS_5376_B1_C12.T7A RPO
gb_gss8:AA054993	-	40.00	141.03	43.32	571	! AA054993 CIT-HSP-2338B21.TR CIT
gb_est9:AA124047	+	39.00	144.58	27.47	247	! AA124047 m22b11.r1 Barstead m
gb_est34:AA178735	+	39.00	142.83	34.41	301	! AA178735 u331b10.y1 Sugano mous
gb_est6:WA14823	+	39.00	138.96	56.47	465	! WA14823 mb3d10.r1 Soares mous
gb_est37:AA1998409	+	39.00	138.59	59.24	485	! AA1998409 701667055.A.thallana,
gb_est11:AA269637	+	39.00	137.74	66.11	534	! AA269637 vs61c10.r1 Soares mous
gb_est36:AA1892890	+	39.00	137.39	69.08	555	! AA1892890 m22b11.y1 Barstead m
gb_gss12:AA0349565	+	38.00	134.63	98.48	489	! AA0349565 RPI11-108N12.TJB RPI
gb_est17:AA617949	+	37.00	134.18	104.34	332	! AA617949 n202405.s1 NCI.CGAP.L
gb_est16:CA26511	+	37.00	131.28	151.28	460	! CA26511 C26511 Rice callus cDNA
gb_est36:AA1895746	+	37.00	131.26	151.66	461	! AA1895746 EST265189 tomato callu
gb_gss13:AA0440876	-	37.00	130.52	166.73	501	! AA0440876 HS_5098_B2_B04.T7A RPO
gb_gss8:AA023777	-	37.00	129.19	197.77	582	! AA023777 CGR0157A Cyptosporid
gb_est36:AA1894896	+	37.00	129.09	200.48	589	! AA1894896 EST264339 tomato callu
gb_est11:Z83985	+	36.00	132.95	122.13	246	! Z83985 SZ83985 Porcine small i
gb_est35:AA1842495	+	36.00	132.74	125.52	259	! AA1842495 UI-M-AM1-afy-h-08-0-UI
gb_est31:AA1708671	+	36.00	132.49	129.50	252	! AA1708671 as39b04.x1 Barstead ac
gb_est21:AA952680	+	36.00	132.16	135.22	269	! AA952680 TENS1864.T.cruzi epin
gb_est24:AA197788	+	36.00	131.31	150.78	296	! AA197788 q149c04.x1 NCI.CGAP.B
gb_est23:AA158411	-	36.00	129.97	178.93	344	! AA158411 u24e07.r1 Soares 2NbM
gb_est24:AA183475	+	36.00	129.77	183.68	352	! AA183475 q54b03.x1 Soares.feta
gb_est9:AA125323	+	36.00	129.04	201.62	382	! AA125323 mp80c08.r1 Soares 2NbM
gb_est30:AA1666560	+	36.00	128.79	208.25	393	! AA1666560 mu1408.x1 Soares 2NbM
gb_est15:AA502026	+	36.00	128.72	210.06	396	! AA502026 nh56g12.s1 NCI.CGAP.P
gb_est19:AA764120	-	36.00	128.70	210.66	397	! AA764120 v45109.x1 Soares 2NbM
gb_est30:AA1642667	+	36.00	128.68	211.27	398	! AA1642667 v501f03.r1 Soares mous
gb_gss13:AA044169	+	36.00	128.61	213.08	401	! AA044169 GSWC0231 Trypanosoma
gb_gss6:AA0835983	+	36.00	128.52	215.50	405	! AA0835983 HS_5350_A1_E03.T7A RPO
gb_gss10:AA021716	+	36.00	128.48	216.72	407	! AA021716 HS_2139_A2_D07.T7 CIT
gb_est37:AA1957215	+	36.00	128.11	227.06	424	! AA1957215 u177a10.x1 Sugano mous

gb\_gss13:AA0431584 - 36.00 127.93 232.56 433 ! AQ431584 HS\_5093\_A2\_C08.SP6E  
gb\_est25:AA1316043 + 36.00 127.87 234.39 436 ! AI316043 u161f12.y1 Sugano m  
gb\_est11:AA199357 - 36.00 127.83 235.62 438 ! AA199357 mu14h06.r1 Soares 2

seq\_name: gb\_est38:AL036690

seq\_documentation\_block:

LOCUS AL036690 171 bp mRNA EST 27-SEP-1999  
DEFINITION DKF2P564D2463\_r1 564 (synonym: hfr2) Homo sapiens CDNA clone  
DKF2P564D2463 5', mRNA sequence.

ACCESSION AL036690

VERSION AL036690.3 GI:5927859

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 171)

AUTHORS Duesterhoeft,A., Lauber,J., Mewes,H.W., Gassenhuber,J. and

Wiemann,S.

TITLE EST (Duesterhoeft, et al.)

JOURNAL Unpublished (1999)

COMMENT On Jul 7, 1999 this sequence version replaced gi:5866258.

Contact: Duesterhoeft A

MIPS

Am Klopferspitz 18a D-82152 Martinsried, Germany

This is the 5' sequence of the clone insert

Clone from S. Wiemann, Molecular Genome Analysis, German Cancer

Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;

sequenced by Qiagen within the CDNA sequencing consortium of the

German Genome Project.

No sl sequence available.

This clone is available at the RZPD in Berlin.

Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059

Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES

source

1..171

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="DKF2P564D2463"

/clone\_11b="564 (synonym: hfr2)"

/tissue\_type="brain"

/dev\_stage="fetal"

/lab\_host="X1-2blue"

/note="vector: pAMP1; Site\_1: NotI; Site\_2: SalI"

BASE COUNT 36 a 53 c 60 g 22 t

ORIGIN

alignment\_scores:

Quality: 50.00 Length: 10

Ratio: 5.000 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-08-653-294-23 x AL036690 ..

Align seg 1/1 to: AL036690 from: 1 to: 171

1 ArgGluAspLeuArgThrLeuLeuArgTyr 10

|||||

62 CGAGAGGACCTGCGGACCTCGCTCGCTAC 91

seq\_name: gb\_gss8:AO039738

seq\_documentation\_block:

LOCUS AO039738 380 bp DNA GSS 11-JUL-1998

DEFINITION CIT-HSP-2317E17.TF CIT-HSP Homo sapiens genomic clone 2317E17,

genomic survey sequence.

ACCESSION AO039738

VERSION AO039738.1 GI:3305570

KEYWORDS GSS.

SOURCE human.

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 380)  
 AUTHORS Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,  
 Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,  
 Simon,M. and Venter,J.C.  
 TITLE Use of a random BAC End Sequence Database for Sequence-Ready Map  
 Building (1998)  
 JOURNAL Unpublished (1998)  
 COMMENT Other\_GSSs: CIT-HSP-2317E17.TR  
 Contact: Mark Adams  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 0200  
 Fax: 301 838 0208  
 Email: mdadams@tigr.org  
 Clones are available from Research Genetics (info@resgen.com). BAC  
 end search page:  
 http://www.tigr.org/cdb/humgen/bac\_end\_search/bac\_end\_search.html.  
 Seq primer: M13-21  
 Class: BAC ends.

## FEATURES

source Location/Qualifiers

1..380  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="2317E17"  
 /clone\_lib="CIT-HSP"  
 /sex="Male"  
 /cell\_type="Sperm"  
 /note="Vector: pBelobAC11; Site\_1: HindIII; Site\_2:  
 HindIII"

BASE COUNT 126 a 79 c 72 g 103 t  
 ORIGIN

## alignment\_scores:

Quality: 44.00 Length: 10  
 Ratio: 4.889 Gaps: 0  
 Percent Similarity: 90.000 Percent Identity: 90.000

## alignment\_block:

US-08-653-294-23 x AQ039738 ..

Align seg 1/1 to: AQ039738 from: 1 to: 380

1 ArgGluAspLeuArgThrLeuLeuArgTyr 10

251 AGAGAAGATCTCAGAACCCCTTTTCATAC 280

seq\_name: gb\_gss13:AQ440598

## seq\_documentation\_block:

LOCUS AQ440598 430 bp DNA 31-MAR-1999  
 DEFINITION HS\_5089\_B1\_C10.SP6E RPCI-11 Human Male BAC Library Homo sapiens  
 genomic clone Plate-665 Col-19 Row-F, genomic survey sequence.  
 ACCESSION AQ440598  
 VERSION AQ440598.1 GI:4551937  
 KEYWORDS GSS.  
 SOURCE human.  
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 430)  
 AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,  
 Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and  
 Hood,L.

TITLE Sequence-tagged connectors: A sequence approach to mapping and  
 scanning the human genome

JOURNAL Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)

MEDLINE 99380589

COMMENT On Feb 19, 1999 this sequence version replaced gi:4146013.

Contact: Mahairas GG, Wallace JC, Hood L  
 High Throughput Sequencing Center  
 University of Washington  
 401 Queen Anne Avenue North, Seattle, WA 98109, USA  
 Tel: (206) 616-3618  
 Fax: (206) 616-3887  
 Email: jwallace@u.washington.edu

Clones are derived from the human BAC library RPCI-11. For BAC  
 library availability, please contact Pieter de Jong  
 (pieter@dejong.med.buffalo.edu). Clones may be purchased from  
 BACPAC Resources ([http://bacpac.med.buffalo.edu/ordering\\_bac.htm](http://bacpac.med.buffalo.edu/ordering_bac.htm))  
 or from Resear h Genetics (info@resgen.com). BAC end Web Server:  
<http://www.htsc.washington.edu>  
 Plate: 665 row: F column: 19  
 Seq primer: SP6  
 Class: BAC ends

High quality sequence stop: 430.

## FEATURES

source

Location/Qualifiers  
 1..430  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="Plate-665 Col-19 Row-F"  
 /clone\_lib="RPCI-11 Human Male BAC Library"  
 /sex="male"  
 /note="Vector: pBACe3.6; Genomic sequence of BAC ends"

BASE COUNT 122 a 81 c 84 g 142 t 1 others  
 ORIGIN

## alignment\_scores:

Quality: 44.00 Length: 10  
 Ratio: 4.889 Gaps: 0  
 Percent Similarity: 90.000 Percent Identity: 90.000

## alignment\_block:

US-08-653-294-23 x AQ440598/rev ..

Align seg 1/1 to reverse of: AQ440598 from: 1 to: 430

1 ArgGluAspLeuArgThrLeuLeuArgTyr 10

177 AGAGAAGATCTCAGAACCCCTTTTCATAC 148

seq\_name: gb\_est20:AA879637

## seq\_documentation\_block:

LOCUS AA879637 347 bp mRNA 26-MAR-1998  
 DEFINITION Vx38b05.r1 Stragatene mouse lung 937302 Mus musculus cDNA clone  
 IMAGE:1277457 5' similar to gb:X03945\_cds1 HLA CLASS I  
 HISTOCOMPATIBILITY ANTIGEN, B-27 ALPHA (HUMAN); gb:M27034 Mouse MHC  
 class I D-region cell surface antigen (MOUSE);, mRNA sequence.

ACCESSION AA879637

VERSION AA879637.1 GI:2990342

## KEYWORDS

EST.

## SOURCE

house mouse.

## ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 347)

AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,  
 Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,  
 Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,  
 Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and  
 Waterston,R.

The WashU-HMI Mouse EST Project

## TITLE

Unpublished (1996)

## JOURNAL

COMMENT

On Jan 9, 1998 this sequence version replaced gi:937414.

Contact: Marra M/Mouse EST Project

WashU-HMI Mouse EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810



IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:581556

Trace considered overall poor quality

Seq primer: -28mi3 rev2 Et from AmerSham

High quality sequence stop: 1.

Location/Qualifiers

#### FEATURES

source

1. .468

/organism="Mus musculus"

/strain="C3H"

/db\_xref="taxon:10090"

/clone="IMAGE:1049980"

/clone\_lib="Barstead mouse myotubes MPLRB5"

/cell\_line="C2C12"

/lab\_host="DH10B"

/note="Vector: pT73D-Pac (Pharmacia) with a modified

polylinker; Site\_1: EcoRI; Site\_2: NotI; 1st strand cDNA

was primed with a Not I - oligo(dT) primer [5'

TGTTACGAATCTGAAGTGGAGCGCCGCTTTTCTTTTCTTTTCTTTT

3']; double-stranded cDNA was ligated to Eco RI adaptors

[AATCGGATCTTG], digested with Not I and cloned into the

Not I and Eco RI sites of the modified pT7T3 vector.

Library constructed by Bob Barstead. The C2C12 cell line

(available from ATCC, catalog # CRL-1772) differentiates

rapidly, forming contractile myotubes and producing

characteristic muscle proteins."

104 a 126 c 169 g 69 t

BASE COUNT

ORIGIN

#### alignment\_scores:

Quality: 43.00 Length: 10  
Ratio: 4.778 Gaps: 0  
Percent Similarity: 90.000 Percent Identity: 90.000

#### alignment\_block:

US-08-653-294-23 x AA592217

Align seg 1/1 to: AA592217 from: 1 to: 468

1 ArgGluspleuarghThrLeuLeuArgTyr 10

|||||

133 CGAGTGGACCTGAGGACCTGCTCCGCTAC 162

seq\_name: gb\_est44:AW208428

seq\_documentation\_block:

LOCUS AW208428 694 bp mRNA EST 03-DEC-1999

DEFINITION u060c03.x1 NCI\_CGAP\_Mam1 Mus musculus cDNA clone IMAGE:2646916 3'

similar to gb:M32320 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-37

B\*3701 ALPHA (HUMAN); gb:M14827 Mouse MHC class I H-2K gene

(MOUSE); mRNA sequence.

ACCESSION AW208428

VERSION AW208428.1 GI:5514422

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 694)

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

On Jun 5, 1998 this sequence version replaced gi:3188321.

Other\_ESTs: u060c03.y1

Contact: Robert Strausberg, Ph.D.

Tel: (301) 496-1550

Email: Robert.Strausberg@nih.gov

Tissue procurement: Gilbert Smith, Ph.D.

CDNA Library Preparation: Life Technologies, Inc.

CNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:  
www-bio.llnl.gov/bbrp/image/image.html

MGI:1027368

Seq primer: -400P from Gibco

High quality sequence stop: 420.

Location/Qualifiers

#### FEATURES

source

1. .694

/organism="Mus musculus"

/strain="FVB/N"

/db\_xref="taxon:10090"

/clone="IMAGE:2646916"

/clone\_lib="NCI\_CGAP Mam1"

/tissue\_type="tumor, biopsy sample"

/dev\_stage="3 months, virgin"

/lab\_host="DH10B"

/note="Organ: mammary; Vector: pCMV-SPORT6; Site\_1: SalI;

Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dT.

Library constructed by Life Technologies. Investigator

providing samples: Gilbert Smith, NIH

BASE COUNT 148 a 206 c 228 g 111 t 1 others

ORIGIN

#### alignment\_scores:

Quality: 43.00 Length: 10  
Ratio: 4.778 Gaps: 0  
Percent Similarity: 90.000 Percent Identity: 90.000

#### alignment\_block:

US-08-653-294-23 x AW208428

Align seg 1/1 to: AW208428 from: 1 to: 694

1 ArgGluspleuarghThrLeuLeuArgTyr 10

|||||

231 CGAGTGGACCTGAGGACCTGCTCCGCTAC 260

seq\_name: gb\_est29:AI556868

seq\_documentation\_block:

LOCUS AI556868 249 bp mRNA EST 05-JUL-1999

DEFINITION UI-R-C2p-ro-g-11-0-UI-s1 UI-R-C2p Rattus norvegicus cDNA clone

UI-R-C2p-ro-g-11-0-UI 3', mRNA sequence.

ACCESSION AI556868

VERSION AI556868.1 GI:4489231

KEYWORDS EST.

SOURCE Norway rat.

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 249)

Bonaldo, M.F., Lennon, G. and Soares, M.B.

Normalization and subtraction: two approaches to facilitate gene

discovery

Genome Res. 6 (9), 791-806 (1996)

On May 18, 1998 this sequence version replaced gi:3138390.

Contact: Soares, MB

Program for Rat Gene Discovery and Mapping

University of Iowa

451 Eckstein Medical

Tel: 319 335 8250

Fax: 319 335 9565

Email: msoares@blue.weeg.uiowa.edu

The sequence tag present in the cDNA between the NotI site and the

oligo-dT track served to identify it as a clone from the normalized

adult 12-day-embryo library. cDNA Library Preparation: M. Fatima

Bonaldo, Ph.D. Clone distribution: clones will be available through

Research Genetics (www.resgen.com) This clone is also available

through the I.M.A.G.E. Consortium at LLNL (info@image.llnl.gov).

IMAGE ID=1790074

Seq primer: M13 Forward

**FEATURES**  
**source**  
 POLYA-No. Location/Qualifiers  
 1. 249  
 /organism="Rattus norvegicus"  
 /strain="Sprague-Dawley"  
 /db\_xref="taxon:10116"  
 /clone="UI-R-C2p-ro-d-11-0-UI"  
 /clone.lib="UI-R-C2p"  
 /dev\_stage="adult"  
 /lab\_host="DH10B (Life Technologies)"  
 /note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site\_1: Not I; Site\_2: Eco RI; The UI-R-C2p library is a subtracted library derived from the UI-R-C1 library, which is a subtracted library derived from the UI-R-C0 library. The UI-R-C0 library consisted of a mixture of individually tagged normalized libraries constructed from rat placenta, adult lung, brain, liver, kidney, heart, spleen, ovary, muscle, 8, 12 and 18-day embryo. The tag is a string of 3-5 nucleotides present between the Not I site and the oligo-dT track which allows identification of the library of origin of a clone within the mixture. The subtracted library (UI-R-C2p) was constructed as follows: PCR amplified cDNA inserts from UI-R-C1 clones from which 3' ESRs had been derived was used as a driver in a hybridization with the UI-R-C1 library in the form of single-stranded circles. The remaining single-stranded circles (subtracted library) was purified by hydroxyapatite column chromatography, converted to double-stranded circles and electroporated into DH10B bacteria (Life Technologies) to generate the UI-R-C2p library. This procedure has been previously described (Bonaldo, Lennon and Soares, Genome Research 6: 791-806, 1996)."  
 67 a 31 c 78 g 73 t  
 BASE COUNT  
 ORIGIN

**alignment\_scores:**  
 Quality: 40.00 Length: 8  
 Ratio: 5.000 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

**alignment\_block:**  
 US-08-653-294-23 x AI556868 ..  
 Align seg 1/1 to: AI556868 from: 1 to: 249

3 AspLeuArgThrLeuLeuArgTyr 10  
 |||||  
 169 GACCTACGTACTTTATTAAGATAT 192

**seq\_name:** gb\_est23:AI124815

**seq\_documentation\_block:**  
 LOCUS AI124815 414 bp mRNA EST 11-SEP-1998  
 DEFINITION am56e06.xl Johnston frontal cortex Homo sapiens cDNA clone IMAGE:1539586 3', similar to gb:M24038.cdsl HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, BW-44(B-12) B\*4402 (HUMAN); contains MER22.t3 TAR1 repetitive element ;, mRNA sequence.  
 ACCESSION AI124815  
 VERSION AI124815.1 GI:3593329  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 414)  
 Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S., Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, K., Wyllie, T., Waterston, R. and Wilson, R. WashU-NCI human EST Project  
 Unpublished (1997)

**REFERENCE**  
**AUTHORS**  
 TITLE  
 JOURNAL

**seq\_documentation\_block:**  
 LOCUS AQ798581 477 bp DNA GSS 09-AUG-1999  
 DEFINITION HS\_2238\_B1\_A11\_MR CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=2238 Col=21 Row=B, genomic survey sequence.  
 ACCESSION AQ798581  
 VERSION AQ798581.1 GI:5715913  
 KEYWORDS GSS.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 477)  
 Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T., Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and Hood, L.  
 Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome  
 Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)  
 99380589  
 Contact: Mahairas GG, Wallace JC, Hood L  
 High Throughput Sequencing Center  
 University of Washington  
 401 Queen Anne Avenue North, Seattle, WA 98109, USA  
 Tel: (206) 616-3618  
 Fax: (206) 616-3887

**FEATURES**  
**source**  
 1. 414  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:1539586"  
 /clone.lib="Johnston frontal cortex"  
 /sex="male"  
 /tissue\_type="pooled frontal lobe"  
 /dev\_stage="adult"  
 /lab\_host="SOLR (kanamycin resistant)"  
 /note="Organ: brain; Vector: Bluescript SK-; Site\_1: EcoRI; Stanley Neuropathology Consortium (www.stanleylab.org) brains S-58, S-65, S-67, S-78. Random + oligo-dT primed into EcoRI site of ZAP II Vector. Mass excised. Avg insert length 1.9kb. Custom library provided by Dr. Nancy Johnston [(410) 614-3918, nlj@welchlink.welch.jhu.edu]."  
 80 a 140 c 136 g 58 t  
 BASE COUNT  
 ORIGIN

**alignment\_scores:**  
 Quality: 40.00 Length: 10  
 Ratio: 4.444 Gaps: 0  
 Percent Similarity: 90.000 Percent Identity: 80.000

**alignment\_block:**  
 US-08-653-294-23 x AI124815 ..  
 Align seg 1/1 to: AI124815 from: 1 to: 414

1 ArgGluAspLeuArgThrLeuLeuArgTyr 10  
 |||||  
 284 CGAGAGACCTGGCACCGCGCTCGCTAC 313

**seq\_name:** gb\_gss5:AQ798581

**seq\_documentation\_block:**  
 LOCUS AQ798581 477 bp DNA GSS 09-AUG-1999  
 DEFINITION HS\_2238\_B1\_A11\_MR CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=2238 Col=21 Row=B, genomic survey sequence.  
 ACCESSION AQ798581  
 VERSION AQ798581.1 GI:5715913  
 KEYWORDS GSS.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 477)  
 Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T., Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and Hood, L.  
 Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome  
 Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)  
 99380589  
 Contact: Mahairas GG, Wallace JC, Hood L  
 High Throughput Sequencing Center  
 University of Washington  
 401 Queen Anne Avenue North, Seattle, WA 98109, USA  
 Tel: (206) 616-3618  
 Fax: (206) 616-3887

**REFERENCE**  
**AUTHORS**  
 TITLE  
 JOURNAL

**seq\_documentation\_block:**  
 LOCUS AQ798581 477 bp DNA GSS 09-AUG-1999  
 DEFINITION HS\_2238\_B1\_A11\_MR CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=2238 Col=21 Row=B, genomic survey sequence.  
 ACCESSION AQ798581  
 VERSION AQ798581.1 GI:5715913  
 KEYWORDS GSS.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 477)  
 Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T., Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and Hood, L.  
 Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome  
 Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)  
 99380589  
 Contact: Mahairas GG, Wallace JC, Hood L  
 High Throughput Sequencing Center  
 University of Washington  
 401 Queen Anne Avenue North, Seattle, WA 98109, USA  
 Tel: (206) 616-3618  
 Fax: (206) 616-3887

**COMMENT**  
 On Jan 17, 1998 this sequence version replaced gi:1899887.  
 Contact: Wilson RK  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info:image.llnl.gov) for further information.  
 Seq primer: "40ml3 fwd" Ex from Amersham.

Email: jwallaceu.washington.edu  
 Clones may be purchased from Research Genetics (info@resgen.com).  
 BAC end Web Server: <http://www.htsc.washington.edu>  
 Plate: 2238 row: B column: 21  
 Seq primer: M13 Reverse  
 Class: BAC ends  
 High quality sequence stop: 477.

Location/Qualifiers  
 1..477  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="Plate:2238 Col-21 Row-B"  
 /clone\_lib="CIT Approved Human Genomic Sperm Library D"  
 /sex="male"  
 /note="Organ: sperm; Vector: pBelobAC11; BAC Clones in E-Coli DH10B"

## FEATURES

source

BASE COUNT 131 a 98 c 122 g 122 t 4 others

ORIGIN

alignment\_scores:  
 Quality: 40.00 Length: 10  
 Ratio: 4.444 Gaps: 0  
 Percent Similarity: 90.000 Percent Identity: 80.000

alignment\_block:

US-08-653-294-23 x A0798581/rev ..

Align seg 1/1 to reverse of: A0798581 from: 1 to: 477

1 ArgGluAspLeuArgThrLeuLeuArgTyr 10

|||||

111 AGAGAAGACTTGAGAACCTGCTATCATTC 82

seq\_name: gb\_est25:A1268994

seq\_documentation\_block:

LOCUS A1268994 504 bp mRNA EST 28-JAN-1999  
 DEFINITION g167h01.x1 NCI\_CGAP\_Kid3 Homo sapiens cDNA clone IMAGE:1864561 3', similar to gb:X51804 PUTATIVE RECEPTOR PROTEIN (HUMAN);, mRNA sequence.

ACCESSION A1268994

VERSION A1268994.1 GI:3888161

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 504)

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT On Jun 2, 1997 this sequence version replaced gi:2285386.

Contact: Robert Strausberg, Ph.D.

Tel: (301) 496-1550

Email: Robert.Strausberg@nih.gov

Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: M. Bento Soares, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.

CDNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

[www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)

Insert Length: 594 Std Error: 0.00

Seq primer: -400P from Gibco

High quality sequence stop: 289.

Location/Qualifiers

1..504

/organism="Homo sapiens"

/db\_xref="taxon:9606"

## FEATURES

source

/clone="IMAGE:1864561"  
 /clone\_lib="NCI\_CGAP\_Kid3"  
 /lab\_host="DH10B"  
 /note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(df) primer, double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. mRNA source: 2 pooled kidneys. Library went through one round of normalization. Library constructed by Bento Soares and M. Fatima Bonaldo."  
 BASE COUNT 116 a 132 c 109 g 145 t 2 others  
 ORIGIN

alignment\_scores:  
 Quality: 40.00 Length: 10  
 Ratio: 4.444 Gaps: 0  
 Percent Similarity: 90.000 Percent Identity: 80.000

alignment\_block:

US-08-653-294-23 x A1268994/rev ..

Align seg 1/1 to reverse of: A1268994 from: 1 to: 504

1 ArgGluAspLeuArgThrLeuLeuArgTyr 10

|||||

413 AAAGAAGATTACGACCTATTCCTAT 384

seq\_name: gb\_gss15:AQ600634

seq\_documentation\_block:

LOCUS AQ600634 557 bp DNA GSS 10-JUN-1999  
 DEFINITION HS\_5376\_B1\_C12\_T7A RBC1-11 Human Male BAC Library Homo sapiens genomic clone Plate-952 Col-23 Row-F, genomic survey sequence.

ACCESSION AQ600634

VERSION AQ600634.1 GI:5060628

KEYWORDS GSS.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 557)

AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.

Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome

Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)

99380589

On Feb 19, 1999 this sequence version replaced gi:4146120.

Contact: Mahairas GG, Wallace JC, Hood L

High Throughput Sequencing Center

University of Washington

401 Queen Anne Avenue North, Seattle, WA 98109, USA

Tel: (206) 616-3618

Fax: (206) 616-3887

Email: jwallace@u.washington.edu

Clones are derived from the human BAC library RBC1-11. For BAC

library availability, please contact Pieter de Jong

(pieterdejong.med.buffalo.edu). Clones may be purchased from

BACPAC Resources ([http://bacpac.med.buffalo.edu/ordering\\_bac.htm](http://bacpac.med.buffalo.edu/ordering_bac.htm))

or from Resear h Genetics (info@resgen.com). BAC end Web Server:

<http://www.htsc.washington.edu>

Plate: 952 row: F column: 23

Seq primer: T7

Class: BAC ends

High quality sequence stop: 557.

Location/Qualifiers

1..557

/organism="Homo sapiens"

/db\_xref="taxon:9606"

## FEATURES

source





## alignment\_block:

US-08-653-294-23 x AAL24047

Percent Similarity: 100.000 Percent Identity: 88.889

Align seg 1/1 to: AAL24047 from: 1 to: 247

alignment\_block:  
US-08-653-294-23 x AI787375

1 ArgGluAspLeuArgThrLeuLeuArg 9

Align seg 1/1 to: AI787375 from: 1 to: 301

87 CGGGAGGATCTACGGACCTTCTGCGA 113

1 ArgGluAspLeuArgThrLeuLeuArg 9  
|||||  
82 CGGGAGGATCTACGGACCTTCTGCGA 108

seq\_name: gb\_est34:AI787375

## seq\_documentation\_block:

LOCUS AI787375 301 bp mRNA EST 02-JUL-1999

DEFINITION u31b10.v1 Sugano mouse kidney mkia Mus musculus cDNA clone  
IMAGE:1921531 5' similar to gb:J05633 INTEGRIN BETA-5 SUBUNIT

PRECURSOR (HUMAN);, mRNA sequence.

ACCESSION AI787375

VERSION AI787375.1 GI:5335091

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

## REFERENCE

## AUTHORS

1 (bases 1 to 301)  
Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,  
Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y.,  
Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R.,  
Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R.,  
Waterston, R. and Wilson, R.

The WashU-NCI Mouse EST Project 1999

Unpublished (1999)

On May 9, 1996 this sequence version replaced gi:1132570.

Other ESTs: u31b10.x1

Contact: Marra M/WashU-NCI Mouse EST Project 1999

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LLNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:977823

Seq primer: custom primer used.

Location/Qualifiers

## FEATURES

## source

1..301  
/organism="Mus musculus"  
/strain="C57BL"  
/db\_xref="taxon:10090"  
/clone="IMAGE:1921531"  
/clone\_lib="Sugano mouse kidney mkia"  
/sex="female"  
/dev\_stage="adult"  
/lab\_host="DH10B"/note="Organ: kidney; Vector: pME18S-FL3; Site\_1: DraIII  
(CACTGTGTG); Site\_2: DraIII (CACCATGTG); 1st strand cDNA  
was primed with an oligo(dT) primer  
[ATGTGGCCTTTTCTTTTCTTTT]; double-stranded cDNA was  
ligated to a DraIII adaptor [TGTGGCCTACTGG], digested  
and cloned into distinct draIII sites of the pME18S-FL3  
vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should  
be used to isolate the cDNA insert. Size selection was  
performed to exclude fragments <1.5kb. Library  
constructed by Dr. Sumio Sugano (University of Tokyo  
Institute of Medical Science). Custom primers for  
sequencing: 5' end primer CTCTGCTCTAAAGCTGCG and 3' end  
primer CGACCTGCAGCTGACACA."

72 a 72 c 98 g 59 t

BASE COUNT  
ORIGIN

## alignment\_scores:

Quality: 39.00  
Ratio: 4.333Length: 9  
Gaps: 0

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2000, 04:05:42 ; Search time 133.56 Seconds  
(without alignments)  
1.773 Million cell updates/sec

Title: US-08-653-294-24  
Perfect score: 51  
Sequence: 1 RESLNRLRGY 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 188963 seqs, 23686106 residues

Total number of hits satisfying chosen parameters: 188963

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database : A\_Geneseq\_36.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query %	Match	Length	DB	ID	Description
1	51	100.0	10	1	R41209	1	Peptide fragment o
2	51	100.0	10	1	R83061	10	HLA-B7 CTL modul
3	51	100.0	10	1	W07515	10	T-cell modulating
4	51	100.0	10	1	W33786	10	Peptide B7.75-84 t
5	51	100.0	10	1	W33796	10	Peptide B7.75-84 t
6	51	100.0	20	1	R92913	20	HLA-B7 CTL modul
7	51	100.0	20	1	R95415	20	HLA-B7.84-75-84 Pa
8	51	100.0	20	1	W33790	20	Peptide B7.84-75/7
9	51	100.0	20	1	W33797	20	Peptide B7.84-75/7
10	51	100.0	25	1	R41207	25	Peptide fragment o
11	51	100.0	25	1	R83073	25	HLA-Bw62 CTL modul
12	51	100.0	25	1	R95431	25	HLA-B7.60-84. Comp
13	51	100.0	25	1	R95419	25	HLA-Bw62.60-84. CO
14	51	100.0	362	1	R12464	362	HLA-B35 antigen. H
15	44	86.3	25	1	R41206	25	Peptide fragment o
16	44	86.3	25	1	R83072	25	HLA-Bw46 CTL modul
17	44	86.3	25	1	R95418	25	HLA-Bw46.60-84. CO
18	44	86.3	366	1	R12465	366	HLA-C exon Cb-1. H
19	44	86.3	366	1	R12466	366	HLA-C exon Cb-2. H
20	44	86.3	366	1	Y07033	366	Breast cancer asso
21	41	80.4	274	1	P80911	274	Consensus sequence
22	37	72.5	10	1	W07518	10	T-cell modulating
23	35	68.6	22	1	R44112	22	HLA epitope. HIV G
24	35	68.6	45	1	R71629	45	HLA-alpha-1. Use o
25	35	68.6	489	1	R47268	489	Partial sequence o
26	35	68.6	489	1	R74305	489	SABP encoded by PC
27	35	68.6	489	1	W04316	489	Tobacco salicylic
28	33	64.7	793	1	R85199	793	Avenacinase-like p
29	33	64.7	850	1	W57445	850	A. thermophilum th
30	33	64.7	1130	1	W05178	1130	H-lats large tumou
31	33	64.7	1201	1	W90345	1201	Drosophila sp. Cos
32	32	62.7	9	1	R42577	9	HLA epitope. HIV G
33	32	62.7	425	1	R11329	425	Alpha subunit of M
34	32	62.7	624	1	R77674	624	Glucosylase from

35 32 62.7 685 1 W31274 Mouse frizzled-8 p  
36 32 62.7 785 1 R04907 Type I interferon  
37 32 62.7 854 1 P70347 Varicella-zoster v  
38 32 62.7 865 1 R20670 Lipoxigenase. Prod  
39 32 62.7 865 1 R23797 Rice lipoxigenase.  
40 32 62.7 868 1 R22615 Varicella-zoster v  
41 31 60.8 10 1 R41211 Peptide fragment o  
42 31 60.8 10 1 W07521 T-cell modulating  
43 31 60.8 93 1 W30133 ZAP-70 N-terminal  
44 31 60.8 212 1 R54922 CdiI polypeptide.  
45 31 60.8 286 1 W54071 Banana bunchy top

## ALIGNMENTS

### RESULT 1

R41209 ID R41209 standard; peptide; 10 AA.  
AC R41209;  
DT 15-MAR-1994 (first entry)  
DE Peptide fragment of Class I HLA peptide.  
KW Human leukocyte antigen; HLA; peptide; transplantation; neoplasia;  
KW Parasitic disease; cytotoxic T lymphocyte; modulation.  
OS Synthetic.  
PN W09317699-A.  
PD 16-SEP-1993.  
PF 25-FEB-1993; U01758.  
PR 02-MAR-1992; US-844716.  
PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
PI Clayberger CA, Krensky AM;  
DR WPI: 93-303134/38.  
PT New peptide(s) based on Class I HLA antigen domains - used for  
PT modulating cytotoxic T-lymphocyte activity towards targets  
PS Claim 11; Page 54; 61pp; English.  
CC The peptide is used to modulate cytotoxic T-lymphocyte (CTL)  
CC activity, either by inhibition or stimulation. It can be used  
CC for inhibiting CTL toxicity in transplantations, for inducing CTL  
CC activity in parasitic diseases and neoplasia and in studies on viral  
CC infection. The peptide can also be used for identifying CTLs which  
CC bind to it and removing subsets of CTLs from a T-cell composition.  
CC This peptide sequence is more commonly found within larger peptide  
CC compounds of not more than 30 amino acids in length.  
SQ Sequence 10 AA:

Query Match 100.0%; Score 51; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. NO. 0.00074;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 RESLNRLRGY 10  
| | | | | | | | | |  
Db 1 RESLNRLRGY 10

### RESULT 2

R83061 ID R83061 standard; peptide; 10 AA.  
AC R83061;  
DT 15-MAY-1996 (first entry)  
DE HLA-B7 CTL modulating peptide (B7.75-84).  
KW Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;  
KW Immunosuppressant; graft versus host disorder; transplantation; therapy;  
KW Class I MHC; HLA-B7.  
OS Synthetic.  
PN W09526979-A1.  
PD 12-OCT-1995.  
PF 05-APR-1995; U04349.  
PR 05-APR-1994; US-222851.  
PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
PI Clayberger C, Krensky AM, Farham P;  
DR WPI: 95-358582/46.  
PT Extension of acceptance period of transplants from MHC unmatched

PT donor hosts - using Class I B75-84 MHC antigen of the recipient  
 PT host  
 PS Claim 13: Page 66: 80pp; English.  
 CC This sequence represents a fragment of a class I major histocompatibility  
 CC complex (MHC) antigen. This sequence corresponds to residues 75-84 of  
 CC the alpha-1 domain of the class I MHC HLA-B7. This sequence, and the  
 CC peptide fragments represented by R83062-R83085, R83090-R83096 and  
 CC R82907-R82913 can be used to extend the period of acceptance by a  
 CC recipient of a transplant from an MHC unmatched donor. The peptides are  
 CC administered to a patient in conjunction with a subtherapeutic amount of  
 CC an immunosuppressant. This is administered to the patient for a limited  
 CC period of time (compared to the lifetime administration for current  
 CC treatments). The peptides particularly modulate (or inhibit) the  
 CC activity of the cytotoxic T lymphocytes (CTLs) of the patient.  
 SQ Sequence 10 AA;

Query Match 100.0%; Score 51; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.00074;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RESLNLRGY 10  
 DB 1 RESLNLRGY 10  
 |||||

RESULT 3  
 W07515 standard; peptide; 10 AA.  
 AC W07515;  
 DT 04-AUG-1997 (first entry)  
 DE T-cell modulator; autoimmune disease; tissue destruction; alpha-domain;  
 KW mammal; major histocompatibility complex; MHC class I; antigen; perforin;  
 KW insulin-dependent diabetes mellitus; multiple sclerosis; inflammation;  
 KW rheumatoid arthritis; psoriasis; pemphigus vulgaris; Sjogren's disease;  
 KW thyroid disease; Hashimoto's thyroiditis; myasthenia gravis; granzyme;  
 KW autologous target cell; cytokine release; T cell activation; therapy.  
 OS Synthetic.  
 PN W09635443-A1.  
 PD 14-NOV-1996.  
 PF 05-APR-1996; U04710.  
 PR 12-MAY-1995; US-440504.  
 PA (SANG-) SANGSTAT MEDICAL CORP.  
 PI Buelow R.  
 DR WPI; 96-518410/51.  
 PT Treatment of auto-immune disease by admin. of peptide(s) corresp. to  
 PT major histocompatibility complex antigens - esp. for delaying onset  
 PT of clinical symptoms of insulin dependent diabetes by modulating T  
 PT cell mediated attack on target cells  
 PS Claim 7: Page 20: 24pp; English.  
 CC W07512-W07518 represent T-cell modulating peptides that can be used in  
 CC the method of the invention. These sequences are based on a portion of  
 CC the generic peptide corresponding to residues 70-91 of the alpha-1 domain  
 CC of the major histocompatibility complex (MHC) class I antigen (see  
 CC W07510). The method is for affecting the course of an autoimmune disease  
 CC involving T-cell mediated destruction of tissue in mammals. These  
 CC peptides are used especially to treat insulin-dependent diabetes  
 CC mellitus, preferably being administered during the pre-clinical stage to  
 CC delay onset of the disease. Other diseases that can be treated are  
 CC multiple sclerosis, rheumatoid arthritis, psoriasis, pemphigus vulgaris,  
 CC Sjogren's disease, thyroid disease, Hashimoto's thyroiditis, myasthenia  
 CC gravis, etc. The peptides modulate T-cell mediated attack on autologous  
 CC target cells, and may also reduce inflammation, swelling, and release of  
 CC cytokines, perforins, granzymes etc. associated with T cell activation.  
 SQ Sequence 10 AA;

Query Match 100.0%; Score 51; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.00074;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RESLNLRGY 10  
 DB 1 RESLNLRGY 10  
 |||||

DB 1 RESLNLRGY 10  
 |||||

RESULT 4  
 W33786 standard; peptide; 10 AA.  
 AC W33786;  
 DT 19-JUN-1998 (first entry)  
 DE Peptide B7.75-84 tested for immunomodulating activity.  
 KW immunomodulating dimer; immunosuppressant drug; CTL activation;  
 KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;  
 KW rejection.  
 OS Synthetic.  
 OS Homo sapiens.  
 PN W09744351-A1.  
 PD 27-NOV-1997.  
 PF 22-MAY-1997; U08689.  
 PR 24-MAY-1996; US-653294.  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI Buelow R, Clayberger C, Krensky AM;  
 DR WPI; 98-085530/08.  
 PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B  
 PT alpha-1 domain, used for preventing rejection of transplants or  
 PT treating autoimmune diseases  
 PS Example 1: Page 19: 41pp; English.  
 CC Peptides W33784-98 and W33778-9 were assayed for their immunomodulating  
 CC activity. A peptide-type compound or variant is claimed which has  
 CC immunomodulating activity, including the N-terminal acylated and/or  
 CC C-terminal amidated or esterified forms of up to 60 amino acids, where  
 CC the peptide-type compound comprises the formula: A-B, where A, B =  
 CC (R aa76-77L) (aa79-84) or (aa84-79) (laa77-76R); aa76 = E or V; aa77 =  
 CC D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a hydrophobic or  
 CC small amino acid; aa82 = R or L; aa83 = G or R; and aa represents amino  
 CC acid. The sequence in the brackets may optionally be absent or truncated  
 CC at any peptide type bond within the brackets. The compounds comprise  
 CC amino acid sequences related to a Class I HLA-B alpha domain (positions  
 CC 79-84). They can be used to inhibit cytotoxic T-lymphocytes (CTL) from  
 CC undesirably attacking cells in a host or in vitro. They can also be  
 CC used in combination with antigenic peptides or proteins of interest to  
 CC activate CTLs. They can also inhibit the proliferation of T cells in  
 CC response to anti-CD3. The peptide can be used for preventing rejection  
 CC of transplants or for treating autoimmune diseases, e.g. diabetes,  
 CC rheumatoid arthritis and lupus erythematosus. The products can also be  
 CC used for detection and diagnosis.  
 SQ Sequence 10 AA;

Query Match 100.0%; Score 51; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.00074;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RESLNLRGY 10  
 DB 1 RESLNLRGY 10  
 |||||

RESULT 5  
 W33796 standard; peptide; 10 AA.  
 AC W33796;  
 DT 19-JUN-1998 (first entry)  
 DE Peptide B7.75-84 tested for immunomodulating activity.  
 KW immunomodulating dimer; immunosuppressant drug; CTL activation;  
 KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;  
 KW rejection.  
 OS Synthetic.  
 OS Homo sapiens.  
 PN W09744351-A1.  
 PD 27-NOV-1997.  
 PF 22-MAY-1997; U08689.  
 PR 24-MAY-1996; US-653294.  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.

PI Beulow R, Clayberger C, Krensky AM;  
DR WPI: 98-086530/08.  
PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B  
PT alpha-1 domain, used for preventing rejection of transplants or  
PT treating autoimmune diseases  
PS Example 1: Page 19; 41pp; English.  
CC Peptides W3784-98 and W3778-9 were assayed for their immunomodulating  
CC activity. A peptide-type compound or variant is claimed which has  
CC immunomodulating activity, including the N-terminal acylated and/or  
CC C-terminal amidated or esterified forms of up to 60 amino acids, where  
CC the peptide-type compound comprises the formula: A-B, where A, B =  
CC (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or V; aa77 =  
CC D, S or R; aa79 = R or G; aa80 = I or N; aa81, aa84 = a hydrophobic or  
CC small amino acid; aa82 = R or L; aa83 = G or R; and aa represents amino  
CC acid. The sequence in the brackets may optionally be absent or truncated  
CC at any peptide type bond within the brackets. The compounds comprise  
CC amino acid sequences related to a Class I HLA-B alpha domain (positions  
CC 79-84). They can be used to inhibit cytotoxic T-lymphocytes (CTL) from  
CC undesirably attacking cells in a host or in vitro. They can also be  
CC used in combination with antigenic peptides or proteins of interest to  
CC activate CTLs. They can also inhibit the proliferation of T cells in  
CC response to anti-CD3. The peptide can be used for preventing rejection  
CC of transplants or for treating autoimmune diseases, e.g. diabetes,  
CC rheumatoid arthritis and lupus erythematosus. The products can also be  
CC used for detection and diagnosis.  
SQ Sequence 10 AA;

Query Match 100.0%; Score 51; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.00074;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RESLNRLRGY 10  
| | | | | | | | | |  
DB 1 RESLNRLRGY 10

RESULT 6  
ID R92913 standard; peptide; 20 AA.  
AC R92913;  
DT 16-MAY-1996 (first entry)  
DE HLA-B7 CTL modulating peptide (B7.84-75/75-84).  
KW Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;  
KW immunosuppressant; graft versus host disorder; transplantation; therapy;  
KW Class I MHC; HLA-B7.  
OS Synthetic.  
PN W09526979-A1.  
PD 12-OCT-1995.  
PR 05-APR-1995; U04349.  
PR 05-APR-1994; US-222851.  
PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
PI Clayberger C, Krensky AM, Farham P;  
DR WPI: 95-358582/46.  
PT Extension of acceptance period of transplants from MHC unmatched  
PT donor hosts - using Class I B75-84 MHC antigen of the recipient  
PT host

PS Example 15: Page 36; 80pp; English.  
CC R83061-R83085, R83090-R83096 and R92907-R92913 represent fragments of  
CC class I major histocompatibility complex (MHC) antigens. This sequence  
CC is an inverted dimer of residues 75-84 of the alpha-1 domain of the class  
CC I MHC HLA-B7. These sequences can be used to extend the period of  
CC acceptance by a recipient of a transplant from an MHC unmatched donor.  
CC The peptides are administered to a patient in conjunction with a  
CC subtherapeutic amount of an immunosuppressant. This is administered to  
CC the patient for a limited period of time (compared to the lifetime  
CC administration for current treatments). The peptides particularly  
CC modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)  
CC of the patient.  
CC Sequence 20 AA;

Query Match 100.0%; Score 51; DB 1; Length 20;

Best Local Similarity 100.0%; Pred. No. 0.0015;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RESLNRLRGY 10  
| | | | | | | | | |  
DB 11 RESLNRLRGY 20

RESULT 7  
ID R95415 standard; peptide; 20 AA.  
AC R95415;  
DT 12-NOV-1996 (first entry)  
DE HLA-B7.84-75-84 Palindrome.  
KW HLA: p74; alpha1-helix; human-leucocyte-associated antigen; inhibitor;  
KW T-cell lysate; membrane protein; mammal; heat shock protein; Hsc70; APC;  
KW B cell; calcium influx; cytotoxic T lymphocyte; CTL; differentiation;  
KW cytolysis; antigen presenting cell.  
OS Synthetic.  
PN W09513288-A1.  
PD 18-MAY-1995.  
PD 10-NOV-1994; U12985.  
PR 10-NOV-1993; US-150493.  
PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
PI Clayberger C, Krensky AM;  
DR WPI: 95-194027/25.  
PT Compsns. comprising lymphoid surface membrane proteins - which may  
PT inhibit cytolytic activity and differentiation of CTLs.  
PS Example; Page 18; 29pp; English.  
CC R95413, and R95415-R95431 represent palindromes and fragments of  
CC human-leucocyte-associated antigens. This sequence represents the  
CC HLA-B7.84-75/75-84 palindrome. These sequences can be used to isolate  
CC the protein p74 from a T-cell lysate. p74 is a T-cell surface membrane  
CC protein associated with T-cell activation in mammalian T-cells, and is  
CC also immunologically cross reactive with the heat shock protein Hsc70.  
CC p74 is found in a limited number of cell types, but is particularly  
CC expressed on B and T cells. p74 can be isolated by lysis of a suitable  
CC cell with an amphoteric detergent, and then passed through an affinity  
CC column containing a covalently bound HLA-B2702 palindromic peptide.  
CC Compositions comprising the extracellular fragment of p74 combined with  
CC HLA-B2702.60-84 (see R95416), induces calcium influx, and inhibits  
CC cytotoxic T lymphocyte (CTL) differentiation or cytolysis. Candidate  
CC compounds can be screened for their effect on the cytolytic activity of  
CC T-cells, by combining them with the extracellular portion of p74 and  
CC determining the amount of binding between the candidate compound and p74.  
CC Modulation of CTL activity can be inhibited in a cellular composition  
CC containing T-cells and antigen presenting cells (APCs) by adding to the  
CC mix the extracellular portion of p74, in an amount sufficient to compete  
CC with p74 for the binding of the p74 ligand.  
SQ Sequence 20 AA;

Query Match 100.0%; Score 51; DB 1; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.0015;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RESLNRLRGY 10  
| | | | | | | | | |  
DB 11 RESLNRLRGY 20

RESULT 8  
ID W33790 standard; peptide; 20 AA.  
AC W33790;  
DT 19-JUN-1998 (first entry)  
DE Peptide B7.84-75/75-84 tested for immunomodulating activity.  
KW Immunomodulating dimer; immunosuppressant drug; CTL activation;  
KW transplplantation; autoimmune disease; Class I HLA-B alpha-1 domain;  
KW rejection.  
OS Synthetic.  
OS Homo sapiens.  
PN W09744351-A1.

PD 27-NOV-1997.  
PF 22-MAY-1997; U08689.  
PR 24-MAY-1996; US-653294.  
PI (STRD ) UNIV LELAND STANFORD JUNIOR.  
PA Beulow R, Clayberger C, Krensky AM;  
DR WPI: 98-086530/08.  
PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B  
PT alpha-1 domain, used for preventing rejection of transplants or  
PT treating autoimmune diseases  
PS Example 1; Page 19; 41pp; English.  
CC Peptides W33784-98 and W33778-9 were assayed for their immunomodulating  
CC activity. A peptide-type compound or variant is claimed which has  
CC immunomodulating activity, including the N-terminal acylated and/or  
CC C-terminal amidated or esterified forms of up to 60 amino acids, where  
CC the peptide-type compound comprises the formula: A-B, where A, B =  
CC (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or V; aa77 =  
CC D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a hydrophobic or  
CC small amino acid; aa82 = R or L; aa83 = G or R; and aa represents amino  
CC acid. The sequence in the brackets may optionally be absent or truncated  
CC at any peptide type bond within the brackets. The compounds comprise  
CC amino acid sequences related to a Class I HLA-B alpha1 domain (positions  
CC 79-84). They can be used to inhibit cytotoxic T-lymphocytes (CTL) from  
CC undesirably attacking cells in a host or in vitro. They can also be  
CC used in combination with antigenic peptides or proteins of interest to  
CC activate CTLs. They can also inhibit the proliferation of T cells in  
CC response to anti-CD3. The peptide can be used for preventing rejection  
CC of transplants or for treating autoimmune diseases, e.g. diabetes,  
CC rheumatoid arthritis and lupus erythematosus. The products can also be  
CC used for detection and diagnosis.  
SQ Sequence 20 AA;

Query Match 100.0%; Score 51; DB 1; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.0015;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RESLRNLGY 10  
Db 11 RESLRNLGY 20  
|||||

RESULT 9  
W33797  
ID W33797 standard; peptide: 20 AA.  
AC W33797;  
DT 19-JUN-1998 (first entry)  
DE Peptide B7:84-75/75-84 tested for immunomodulating activity.  
KW Immunomodulating dimer; immunosuppressant drug; CTL activation;  
KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;  
KW rejection.  
OS Synthetic.  
OS Homo sapiens.  
PN WO9744351-A1.  
PD 27-NOV-1997.  
PF 22-MAY-1997; U08689.  
PR 24-MAY-1996; US-653294.  
PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
PI Beulow R, Clayberger C, Krensky AM;  
DR WPI: 98-086530/08.  
PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B  
PT alpha-1 domain, used for preventing rejection of transplants or  
PT treating autoimmune diseases  
PS Example 1; Page 19; 41pp; English.  
CC Peptides W33784-98 and W33778-9 were assayed for their immunomodulating  
CC activity. A peptide-type compound or variant is claimed which has  
CC immunomodulating activity, including the N-terminal acylated and/or  
CC C-terminal amidated or esterified forms of up to 60 amino acids, where  
CC the peptide-type compound comprises the formula: A-B, where A, B =  
CC (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or V; aa77 =  
CC D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a hydrophobic or  
CC small amino acid; aa82 = R or L; aa83 = G or R; and aa represents amino  
CC acid. The sequence in the brackets may optionally be absent or truncated  
CC at any peptide type bond within the brackets. The compounds comprise

CC amino acid sequences related to a Class I HLA-B alpha1 domain (positions  
CC 79-84). They can be used to inhibit cytotoxic T-lymphocytes (CTL) from  
CC undesirably attacking cells in a host or in vitro. They can also be  
CC used in combination with antigenic peptides or proteins of interest to  
CC activate CTLs. They can also inhibit the proliferation of T cells in  
CC response to anti-CD3. The peptide can be used for preventing rejection  
CC of transplants or for treating autoimmune diseases, e.g. diabetes,  
CC rheumatoid arthritis and lupus erythematosus. The products can also be  
CC used for detection and diagnosis.  
SQ Sequence 20 AA;

Query Match 100.0%; Score 51; DB 1; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.0015;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RESLRNLGY 10  
Db 11 RESLRNLGY 20  
|||||

## RESULT 10

R41207

ID R41207 standard; peptide: 25 AA.

AC R41207;

DT 15-MAR-1994 (first entry)

DE Peptide fragment of Class I HLA peptide.

KW Human leukocyte antigen; HLA; peptide; transplantation; neoplasia;

KW parasitic disease; cytotoxic T lymphocyte; modulation.

OS Synthetic.

PN WO9317699-A.

PD 16-SEP-1993.

PF 25-FEB-1993; U01758.

PR 02-MAR-1992; US-844716.

PA (STRD ) UNIV LELAND STANFORD JUNIOR.

PI Clayberger CA, Krensky AM;

DR WPI: 93-303134/38.

PT New peptide(s) based on Class I HLA antigen domains - used for

PT modulating cytotoxic T-lymphocyte activity towards targets

PS Claim 10; Page 54; 61pp; English.

CC The peptide (or a fragment of at least 10 amino acids, joined at at

CC least one terminus to a sequence other than that of wild type HLA

CC antigen) is used to modulate cytotoxic T-lymphocyte (CTL) activity,

CC either by inhibition or stimulation. It can be used for

CC inhibiting CTL toxicity in transplantations, for inducing CTL

CC activity in parasitic diseases and neoplasia and in studies on viral

CC infection. The peptide can also be used for identifying CTLs which

CC bind to it and removing subsets of CTLs from a T-cell composition.

SQ Sequence 25 AA;

Query Match 100.0%; Score 51; DB 1; Length 25;

Best Local Similarity 100.0%; Pred. No. 0.0019;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RESLRNLGY 10  
Db 16 RESLRNLGY 25  
|||||

## RESULT 11

R83073

ID R83073 standard; peptide: 25 AA.

AC R83073;

DT 16-MAY-1996 (first entry)

DE HLA-Bw62 CTL modulating peptide (Bw62.60-84).

KW Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;

KW immunosuppressant; graft versus host disorder; transplantation; therapy;

KW class I MHC; HLA-Bw62.

OS Synthetic.

PN WO9526979-A1.

PD 12-OCT-1995.

PF 05-APR-1995; U04349.

PR 05-APR-1994; US-222851.  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI Clayberger C, Krensky AM, Parham P;  
 DI WPI: 95-358582/46.  
 PT Extension of acceptance period of transplants from MHC unmatched  
 PT donor hosts - using Class I B75-84 MHC antigen of the recipient  
 PT host  
 PS Example 13; Page 32; 80pp; English.  
 CC R83061-R83085, R83090-R83096 and R92907-R92913 represent fragments of  
 CC class I major histocompatibility complex (MHC) antigens. This sequence  
 CC corresponds to residues 60-84 of the alpha-1 domain of the class I MHC  
 CC HLA-B\*62. These sequences can be used to extend the period of acceptance  
 CC by a recipient of a transplant from an MHC unmatched donor. The peptides  
 CC are administered to a patient in conjunction with a subtherapeutic amount  
 CC of an immunosuppressant. This is administered to the patient for a  
 CC limited period of time (compared to the lifetime administration for  
 CC current treatments). The peptides particularly modulate (or inhibit) the  
 CC activity of the cytotoxic T lymphocytes (CTLs) of the patient.  
 SQ Sequence 25 AA;

Query Match 100.0%; Score 51; DB 1; Length 25;  
 Best Local Similarity 100.0%; Pred. No. 0.0019;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RESLRNLRGY 10  
 DB 16 RESLRNLRGY 25  
 |||||

## RESULT 12

ID R95431 standard; peptide; 25 AA.  
 AC R95431;  
 DT 12-NOV-1996 (first entry)  
 DE HLA-B\*62.60-84.  
 KW HLA; p74; alpha1-helix; human-leucocyte-associated antigen; inhibitor;  
 KW T-cell lysate; membrane protein; mammal; heat shock protein; Hsc70; APC;  
 KW B cell; calcium influx; cytotoxic T lymphocyte; CTL; differentiation;  
 KW cytolysis; antigen presenting cell.  
 OS Synthetic.  
 PN W09513288-A1.  
 PD 18-MAY-1995.  
 PF 10-NOV-1994; U12985.  
 PR 10-NOV-1993; US-150493.  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI Clayberger C, Krensky AM;  
 DI WPI: 95-194027/25.  
 PT Compsns. comprising lymphoid surface membrane proteins - which may  
 PT inhibit cytolytic activity and differentiation of CTLs.  
 PS Example; Page 12; 29pp; English.  
 CC R95413, and R95415-R95431 represent palindromes and fragments of  
 CC human-leucocyte-associated antigens. This sequence represents the  
 CC HLA-B\*62.60-84. These sequences can be used to isolate the protein p74  
 CC from a T-cell lysate. p74 is a T-cell surface membrane protein  
 CC associated with T-cell activation in mammalian T-cells, and is also  
 CC immunologically cross reactive with the heat shock protein Hsc70. p74 is  
 CC found in a limited number of cell types, but is particularly expressed on  
 CC B and T cells. p74 can be isolated by lysis of a suitable cell with an  
 CC amphoteric detergent, and then passed through an affinity column  
 CC containing a covalently bound HLA-B2702 palindromic peptide.  
 CC Compositions comprising the extracellular fragment of p74 combined with  
 CC HLA-B2702 60-84 (see R95416), induces calcium influx, and inhibits  
 CC cytotoxic T lymphocyte (CTL) differentiation or cytolysis. Candidate  
 CC compounds can be screened for their effect on the cytolytic activity of  
 CC T-cells, by combining them with the extracellular portion of p74 and  
 CC determining the amount of binding between the candidate compound and p74.  
 CC Modulation of CTL activity can be inhibited in a cellular composition  
 CC containing T-cells and antigen presenting cells (APCs), by adding to the  
 CC mix the extracellular portion of p74, in an amount sufficient to compete  
 CC with p74 for the binding of the p74 ligand.

SQ Sequence 25 AA;

Query Match 100.0%; Score 51; DB 1; Length 25;  
 Best Local Similarity 100.0%; Pred. No. 0.0019;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RESLRNLRGY 10  
 DB 16 RESLRNLRGY 25  
 |||||

## RESULT 13

ID R95419 standard; peptide; 25 AA.  
 AC R95419;  
 DT 12-NOV-1996 (first entry)  
 DE HLA-B\*62.60-84.  
 KW HLA; p74; alpha1-helix; human-leucocyte-associated antigen; inhibitor;  
 KW T-cell lysate; membrane protein; mammal; heat shock protein; Hsc70; APC;  
 KW B cell; calcium influx; cytotoxic T lymphocyte; CTL; differentiation;  
 KW cytolysis; antigen presenting cell.  
 OS Synthetic.  
 PN W09513288-A1.  
 PD 18-MAY-1995.  
 PF 10-NOV-1994; U12985.  
 PR 10-NOV-1993; US-150493.  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI Clayberger C, Krensky AM;  
 DI WPI: 95-194027/25.  
 PT Compsns. comprising lymphoid surface membrane proteins - which may  
 PT inhibit cytolytic activity and differentiation of CTLs.  
 PS Example; Page 9; 29pp; English.  
 CC R95413, and R95415-R95431 represent palindromes and fragments of  
 CC human-leucocyte-associated antigens. This sequence represents the  
 CC HLA-B\*62.60-84. These sequences can be used to isolate the protein p74  
 CC from a T-cell lysate. p74 is a T-cell surface membrane protein  
 CC associated with T-cell activation in mammalian T-cells, and is also  
 CC immunologically cross reactive with the heat shock protein Hsc70. p74 is  
 CC found in a limited number of cell types, but is particularly expressed on  
 CC B and T cells. p74 can be isolated by lysis of a suitable cell with an  
 CC amphoteric detergent, and then passed through an affinity column  
 CC containing a covalently bound HLA-B2702 palindromic peptide.  
 CC Compositions comprising the extracellular fragment of p74 combined with  
 CC HLA-B2702 60-84 (see R95416), induces calcium influx, and inhibits  
 CC cytotoxic T lymphocyte (CTL) differentiation or cytolysis. Candidate  
 CC compounds can be screened for their effect on the cytolytic activity of  
 CC T-cells, by combining them with the extracellular portion of p74 and  
 CC determining the amount of binding between the candidate compound and p74.  
 CC Modulation of CTL activity can be inhibited in a cellular composition  
 CC containing T-cells and antigen presenting cells (APCs), by adding to the  
 CC mix the extracellular portion of p74, in an amount sufficient to compete  
 CC with p74 for the binding of the p74 ligand.

SQ Sequence 25 AA;

Query Match 100.0%; Score 51; DB 1; Length 25;  
 Best Local Similarity 100.0%; Pred. No. 0.0019;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RESLRNLRGY 10  
 DB 16 RESLRNLRGY 25  
 |||||

## RESULT 14

ID R12464 standard; Protein; 362 AA.  
 AC R12464;  
 DT 29-AUG-1991 (first entry)  
 DE HLA-B\*35 antigen.  
 KW Human leukocyte antigen; probe; major histocompatibility complex;  
 KW MHC; class I.  
 OS Homo sapiens.  
 PN J03112486-A.

PD 14-MAY-1991.  
 PF 22-SEP-1989; 247697.  
 PR 22-SEP-1989; JP-247697.  
 PA (OLYU ) OLYMPUS OPTICAL KK.  
 DR WPI; 91-182991/25.  
 DR N-PSDB; Q12115.  
 PT HLA-B35 gene - used in DNA probe and transformant cells for  
 PT immunising animals, for developing monoclonal antibody.  
 PS Claim 1; Page 1; 11pp; Japanese  
 CC Probes comprising part of the sequence encoding this sequence can  
 CC be used to identify Class I genes. The DNA can be expressed for  
 CC immunisation of animals and prodn. of monoclonal antibodies  
 CC specific for the HLA-B35 antigen. See also J03112485 and J03112487.  
 SQ Sequence 362 AA;

Query Match 100.0%; Score 51; DB 1; Length 362;  
 Best Local Similarity 100.0%; Pred. No. 0.031;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RESLNLRGY 10  
 | | | | | | | |  
 Db 99 RESLNLRGY 108

RESULT 15  
 R41206  
 ID R41206 standard; peptide; 25 AA.  
 AC R41206;  
 DT 15-MAR-1994 (first entry)  
 DE Peptide fragment of Class I HLA peptide.  
 KW Human leukocyte antigen; HLA; peptide; transplantation; neoplasia;  
 KW parasitic disease; cytotoxic T lymphocyte; modulation.  
 OS Synthetic.  
 PN WO9317699-A.  
 PD 16-SEP-1993.  
 PF 25-FEB-1993; U01758.  
 PR 02-MAR-1992; US-844716.  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI Clayberger CA, Krensky AM;  
 DR WPI; 93-303134/38.  
 PT New peptide(s) based on Class I HLA antigen domains - used for  
 PT modulating cytotoxic T-lymphocyte activity towards targets  
 PS Claim 9; Page 53; 61pp; English.  
 CC The peptide (or a fragment of at least 10 amino acids, joined at at  
 CC least one terminus to a sequence other than that of wild type HLA  
 CC antigen) is used to modulate cytotoxic T-lymphocyte (CTL) activity,  
 CC either by inhibition or stimulation. It can be used for  
 CC inhibiting CTL toxicity in transplantations, for inducing CTL  
 CC activity in parasitic diseases and neoplasia and in studies on viral  
 CC infection. The peptide can also be used for identifying CTLs which  
 CC bind to it and removing subsets of CTLs from a T-cell composition.  
 SQ Sequence 25 AA;

Query Match 86.3%; Score 44; DB 1; Length 25;  
 Best Local Similarity 90.0%; Pred. No. 0.04;  
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RESLNLRGY 10  
 | | | | | | | |  
 Db 16 RVSLNLRGY 25

Search completed: February 8, 2000, 04:05:42  
 Job time: 9359 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 7, 2000, 18:04:37 ; Search time 111.22 Seconds  
(without alignments)  
4.241 Million cell updates/sec

Title: US-08-653-294-24

Perfect score: 51

Sequence: 1 RESLNRLRGY 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 142080 seqs, 47169319 residues

Total number of hits satisfying chosen parameters: 142080

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : PIR\_62:\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	51	100.0	137	2	I38875
2	51	100.0	137	2	I38876
3	51	100.0	137	2	I38860
4	51	100.0	137	2	I38874
5	51	100.0	181	2	I59188
6	51	100.0	270	1	HLH040
7	51	100.0	274	2	I68774
8	51	100.0	274	2	S24439
9	51	100.0	300	2	I68701
10	51	100.0	308	2	I36956
11	51	100.0	350	2	I68747
12	51	100.0	350	2	I54308
13	51	100.0	354	2	S24436
14	51	100.0	354	2	S24437
15	51	100.0	354	2	S24438
16	51	100.0	354	2	S24440
17	51	100.0	354	2	S24433
18	51	100.0	358	2	S03538
19	51	100.0	361	2	I54418
20	51	100.0	362	1	HLHUB7
21	51	100.0	362	2	S77966
22	51	100.0	362	2	G01230
23	51	100.0	362	2	S16789
24	51	100.0	362	2	A45880
25	51	100.0	362	2	I37519
26	51	100.0	362	2	I37520
27	51	100.0	362	2	I37522
28	51	100.0	362	2	I84431
29	51	100.0	362	2	I72755
30	51	100.0	362	2	I84488

#### ALIGNMENTS

##### RESULT 1

I38875

MHC class I antigen - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 07-Jun-1996 #sequence\_revision 07-Jun-1996 #text\_change 23-Jul-1999

C:Accession: I38875

R:Garber, T.L.; Butler, L.M.; Trachtenberg, E.A.; Erlich, H.A.; Rickards, O.; De Stef

Immunogenetics 42, 19-27, 1995

A:Title: HLA-B alleles of the Cayapa of Ecuador: new B39 and B15 alleles.

A:Reference number: I38860; MUID:95317819

A:Accession: I38875

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-137 <RES>

A:Cross-references: EMBL:U15639; NID:g930332; PIDN:AAA74046.1; PID:g930333

C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 100.0%; Score 51; DB 2; Length 137;

Best Local Similarity 100.0%; Pred. No. 0.0045;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RESLNRLRGY 10

Db 40 RESLNRLRGY 49

##### RESULT 2

I38876

MHC class I antigen - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 07-Jun-1996 #sequence\_revision 07-Jun-1996 #text\_change 23-Jul-1999

C:Accession: I38876

R:Garber, T.L.; Butler, L.M.; Trachtenberg, E.A.; Erlich, H.A.; Rickards, O.; De Stef

Immunogenetics 42, 19-27, 1995

A:Title: HLA-B alleles of the Cayapa of Ecuador: new B39 and B15 alleles.

A:Reference number: I38860; MUID:95317819

A:Accession: I38876

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-137 <RES>

A:Cross-references: EMBL:U15640; NID:g930334; PIDN:AAA74047.1; PID:g930335

C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 100.0%; Score 51; DB 2; Length 137;

Best Local Similarity 100.0%; Pred. No. 0.0045;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RESLNRLRGY 10

Db 40 RESLNRLRGY 49

HLA-B\*5501 - human  
HLA-B\*5502 - human  
HLA-B\*5601 - human  
MHC class I histoc  
HLA-B alpha-chain  
lymphocyte antigen  
HLA-B\*5401 - human  
MHC class I protei  
HLA-B35 variant -  
lymphocyte antigen  
lymphocyte antigen  
MHC HLA-Bw42, HLA-  
MHC HLA-B14 chain  
MHC HLA-B18 chain

31 51 100.0 362 2 I72752  
32 51 100.0 362 2 I72753  
33 51 100.0 362 2 I72754  
34 51 100.0 362 2 I38437  
35 51 100.0 362 2 I37492  
36 51 100.0 362 2 I54505  
37 51 100.0 362 2 I56130  
38 51 100.0 362 2 I36962  
39 51 100.0 362 2 I37475  
40 51 100.0 362 2 I56149  
41 51 100.0 362 2 I59655  
42 51 100.0 362 2 I59651  
43 51 100.0 362 2 I61865  
44 51 100.0 362 2 I61859  
45 51 100.0 362 2 I61860

```
RESULT 3
I38860
MHC class I antigen - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change 23-Jul-1999
C:Accession: I38860
R:Garber, T.L.; Butler, L.M.; Trachtenberg, E.A.; Erlich, H.A.; Rickards, O.; De Stefani
Immunogenetics 42, 19-27, 1995
A:Title: HLA-B alleles of the Cayapa of Ecuador: new B39 and B15 alleles.
A:Reference number: I38860; MUID:95317819
A:Accession: I38860
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-137 <RES>
A:Cross-references: EMBL:U14756; NID:g930328; PIDN:AA50171.1; PID:g930329
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 100.0%; Score 51; DB 2; Length 137;
Best Local Similarity 100.0%; Pred. No. 0.0045;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RESLNRLRGY 10
|||||
Db 40 RESLNRLRGY 49

RESULT 4
I38874
MHC class I antigen - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change 23-Jul-1999
C:Accession: I38874
R:Garber, T.L.; Butler, L.M.; Trachtenberg, E.A.; Erlich, H.A.; Rickards, O.; De Stefani
Immunogenetics 42, 19-27, 1995
A:Title: HLA-B alleles of the Cayapa of Ecuador: new B39 and B15 alleles.
A:Reference number: I38860; MUID:95317819
A:Accession: I38874
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-137 <RES>
A:Cross-references: EMBL:U15638; NID:g930330; PIDN:AAA74045.1; PID:g930331
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 100.0%; Score 51; DB 2; Length 137;
Best Local Similarity 100.0%; Pred. No. 0.0045;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RESLNRLRGY 10
|||||
Db 40 RESLNRLRGY 49

RESULT 5
I59188
MHC cell surface glycoprotein - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 23-Jul-1999
C:Accession: I59188
R:Bronson, S.K.; Pei, J.; Taillon-Miller, P.; Chorney, M.J.; Geraghty, D.E.; Chaplin, D.
Proc. Natl. Acad. Sci. U.S.A. 88, 1676-1680, 1991
A:Title: Isolation and characterization of yeast artificial chromosome clones linking th
A:Reference number: I59188; MUID:91156671
A:Accession: I59188
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-181 <RES>
A:Cross-references: GB:M59841; NID:g187697; PIDN:AAA59623.1; PID:g187698
C:Genetics:
A:Gene: GDB:HLA-B
A:Cross-references: GDB:120048; OMIM:142830
```

```
A:Map position: 6p21.3-6p21.3
A:Introns: 90/1
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology
C:Keywords: glycoprotein

Query Match 100.0%; Score 51; DB 2; Length 181;
Best Local Similarity 100.0%; Pred. No. 0.0061;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RESLNRLRGY 10
|||||
Db 74 RESLNRLRGY 83

RESULT 6
HLH40
MHC class I histocompatibility antigen HLA-B40 alpha chain - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 28-Aug-1985 #sequence_revision 28-Aug-1985 #text_change 02-Sep-1997
C:Accession: A02186
R:Lopez de Castro, J.A.; Bragado, R.; Strong, D.M.; Strominger, J.L.
Biochemistry 22, 3961-3969, 1983
A:Title: Primary structure of papain-solubilized human histocompatibility antigen HLA
A:Reference number: A02186; MUID:84000412
A:Accession: A02186
A:Molecule type: protein
A:Residues: 1-270 <LOP>
C:Comment: Comparison of a number of class I allelic sequences (-B40 vs. -B7 and -A2
y restricted to the alpha-1 and alpha-2 domains. The most conspicuous clustering of v
ossible alloantigenic determinants of these antigens.
C:Genetics:
A:Gene: GDB:HLA-B
A:Cross-references: GDB:120048; OMIM:142830
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology
C:Keywords: duplication; glycoprotein; heterodimer; membrane protein; transplantation
F:1-90/Domain: alpha-1 <EX1>
F:91-181/Domain: alpha-2 <EX2>
F:195-260/Domain: immunoglobulin homology <IMM>
F:86/Binding site: carbohydrate (Asn) (covalent) #status experimental
F:101-163,202-258/disulfide bonds: #status predicted

Query Match 100.0%; Score 51; DB 1; Length 270;
Best Local Similarity 100.0%; Pred. No. 0.0093;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RESLNRLRGY 10
|||||
Db 75 RESLNRLRGY 84

RESULT 7
I68774
MHC HLA-B39 chain - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change 23-Jul-1999
C:Accession: I68774
R:Mueller, C.A.; Engler-Blum, G.; Gekeler, V.; Steiert, I.; Weiss, E.; Schmidt, H.
Immunogenetics 30, 200-207, 1989
A:Title: Genetic and serological heterogeneity of the supertypic HLA-B locus specific
A:Reference number: I54463; MUID:89379286
A:Accession: I68774
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-274 <RES>
A:Cross-references: GB:M29865; NID:g187676; PIDN:AAA36223.1; PID:g187677
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 100.0%; Score 51; DB 2; Length 274;
Best Local Similarity 100.0%; Pred. No. 0.0095;
```

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RESLRNLRGY 10

Db 75 RESLRNLRGY 84

RESULT 8

S24439

Class I histocompatibility antigen HLA-B-3901 - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 22-Nov-1993 #sequence\_revision 01-Nov-1996 #text\_change 07-Nov-1997

C:Accession: S24439

R:Watkins, D.I.; McAdam, S.N.; Liu, X.; Strang, C.R.; Milford, E.L.; Levine, C.G.; Garbe

Nature 357, 329-333, 1992

A:Title: New recombinant HLA-B alleles in a tribe of South American Amerindians indicate

A:Reference number: S24027; MUID:92269956

A:Accession: S24439

A:Molecule type: mRNA

A:Residues: 1-274 <WAT>

C:Genetics:

A:Gene: HLA-B-3901

C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

C:Keywords: glycoprotein

F:1-274/Product: class I histocompatibility antigen HLA-B-3901 #status predicted <WAT>

F:1-90/Domain: alpha-1 #status predicted <EX1>

F:91-182/Domain: alpha-2 #status predicted <EX2>

F:196-261/Domain: immunoglobulin homology <IMM>

F:86/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:101-164,203-259/Disulfide bonds: #status predicted

Query Match

Best Local Similarity 100.0%; Score 51; DB 2; Length 274;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RESLRNLRGY 10

Db 75 RESLRNLRGY 84

RESULT 9

I68701

cell surface antigen - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 07-Jun-1996 #sequence\_revision 07-Jun-1996 #text\_change 23-Jul-1999

C:Accession: I68701

R:Arnot, D.; Lillie, J.W.; Auffray, C.; Kappes, D.; Strominger, J.L.

Immunogenetics 20, 237-252, 1984

A:Title: Inter-locus and intra-allelic polymorphisms of HLA class I antigen gene mRNA.

A:Reference number: I54412; MUID:84287690

A:Accession: I68701

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-300 <RES>

A:Cross-references: GB:M27540; NID:g187733; PIDN:AAA59638.1; PID:g386890

C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

C:Keywords: surface antigen

Query Match

Best Local Similarity 100.0%; Score 51; DB 2; Length 300;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RESLRNLRGY 10

Db 37 RESLRNLRGY 46

RESULT 10

I36956

MHC class II chain - chimpanzee (fragment)

C:Species: Pan troglodytes (chimpanzee)

C>Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 23-Jul-1999

C:Accession: I36956

R:Parham, P.; Lawlor, D.A.; Lomen, C.E.; Ennis, P.D.

J. Immunol. 142, 3937-3950, 1989

A:Title: Diversity and diversification of HLA-A,B,C alleles.

A:Reference number: I36956; MUID:89235215

A:Accession: I36956

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-308 <RES>

A:Cross-references: GB:M24044; NID:g176812; PIDN:AAA53423.1; PID:g176813

C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match

Best Local Similarity 100.0%; Score 51; DB 2; Length 308;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RESLRNLRGY 10

Db 45 RESLRNLRGY 54

RESULT 11

I68747

MHC class I lymphocyte antigen - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 07-Jun-1996 #sequence\_revision 07-Jun-1996 #text\_change 23-Jul-1999

C:Accession: I68747

R:Pohla, H.; Kuon, W.; Tabaczewski, P.; Doerner, C.; Weiss, E.H.

Immunogenetics 29, 297-307, 1989

A:Title: Allelic variation in HLA-B and HLA-C sequences and the evolution of the HLA-

A:Reference number: I54457; MUID:89233295

A:Accession: I68747

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-350 <RES>

A:Cross-references: GB:M28204; NID:g576472; PIDN:AAA53257.1; PID:g576473

C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match

Best Local Similarity 100.0%; Score 51; DB 2; Length 350;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RESLRNLRGY 10

Db 87 RESLRNLRGY 96

RESULT 12

I54308

MHC HLA B71 - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 23-Jul-1999

C:Accession: I54308

R:Rodriguez, S.G.; Johnson, A.H.; Hurley, C.K.

Hum. Immunol. 37, 192-194, 1993

A:Title: Molecular characterization of HLA-B71 from an African American individual.

A:Reference number: I54308; MUID:94064392

A:Accession: I54308

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-350 <RES>

A:Cross-references: GB:L07950; NID:g307236; PIDN:AAA59683.1; PID:g307237

C:Genetics:

A:Gene: GDB:HLA-B

A:Cross-references: GDB:L20048; OMIM:142830

A:Map position: 6p21.3-6p21.3

C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match

Best Local Similarity 100.0%; Score 51; DB 2; Length 350;

Best Local Similarity 100.0%; Pred. No. 0.012;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RESLNRLGY 10  
| | | | | | | | | |  
Db 99 RESLNRLGY 108

## RESULT 13

S24436  
class I histocompatibility antigen HLA-B-3504 precursor - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 22-Nov-1993 #sequence\_revision 18-Jul-1997 #text\_change 07-Nov-1997  
C:Accession: S24436  
R:Watkins, D.I.; McAdam, S.N.; Liu, X.; Strang, C.R.; Milford, E.L.; Levine, C.G.; Garber  
Nature 357, 329-333, 1992  
A:Title: New recombinant HLA-B alleles in a tribe of South American Amerindians indicate  
A:Reference number: S24027; MUID:92269956  
A:Accession: S24436  
A:Molecule type: mRNA  
A:Residues: 1-354 <WAT>  
C:Genetics:  
A:Gene: HLA-B-3504  
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology  
C:Keywords: glycoprotein; transmembrane protein  
F:1-17/Domain: signal sequence #status predicted <SIG>  
F:17-354/Product: class I histocompatibility antigen HLA-B-3504 #status predicted <WAT>  
F:17-299/Domain: extracellular #status predicted <EXT>  
F:107-198/Domain: alpha-2 #status predicted <EX2>  
F:212-277/Domain: immunoglobulin homology <IMM>  
F:300-323/Domain: transmembrane #status predicted <TMN>  
F:324-354/Domain: intracellular #status predicted <INT>  
F:102/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:117-180,219-275/Disulfide bonds: #status predicted

Query Match 100.0%; Score 51; DB 2; Length 354;  
Best Local Similarity 100.0%; Pred. No. 0.012;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RESLNRLGY 10  
| | | | | | | | | |  
Db 91 RESLNRLGY 100

## RESULT 14

S24437  
class I histocompatibility antigen HLA-B-4802 precursor - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 22-Nov-1993 #sequence\_revision 18-Jul-1997 #text\_change 07-Nov-1997  
C:Accession: S24437  
R:Watkins, D.I.; McAdam, S.N.; Liu, X.; Strang, C.R.; Milford, E.L.; Levine, C.G.; Garber  
Nature 357, 329-333, 1992  
A:Title: New recombinant HLA-B alleles in a tribe of South American Amerindians indicate  
A:Reference number: S24027; MUID:92269956  
A:Accession: S24437  
A:Molecule type: mRNA  
A:Residues: 1-354 <WAT>  
C:Genetics:  
A:Gene: HLA-B-4802  
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology  
C:Keywords: glycoprotein; transmembrane protein  
F:1-17/Domain: signal sequence #status predicted <SIG>  
F:17-354/Product: class I histocompatibility antigen HLA-B-4802 #status predicted <WAT>  
F:17-299/Domain: extracellular #status predicted <EXT>  
F:107-198/Domain: alpha-2 #status predicted <EX2>  
F:212-277/Domain: immunoglobulin homology <IMM>  
F:300-323/Domain: transmembrane #status predicted <TMN>  
F:324-354/Domain: intracellular #status predicted <INT>  
F:102/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:117-180,219-275/Disulfide bonds: #status predicted

Query Match 100.0%; Score 51; DB 2; Length 354;  
Best Local Similarity 100.0%; Pred. No. 0.012;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RESLNRLGY 10  
| | | | | | | | | |  
Db 91 RESLNRLGY 100

## RESULT 15

S24438  
class I histocompatibility antigen HLA-B-4801 precursor - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 22-Nov-1993 #sequence\_revision 18-Jul-1997 #text\_change 07-Nov-1997  
C:Accession: S24438  
R:Watkins, D.I.; McAdam, S.N.; Liu, X.; Strang, C.R.; Milford, E.L.; Levine, C.G.; Ga  
Nature 357, 329-333, 1992  
A:Title: New recombinant HLA-B alleles in a tribe of South American Amerindians indic  
A:Reference number: S24027; MUID:92269956  
A:Accession: S24438  
A:Molecule type: mRNA  
A:Residues: 1-354 <WAT>  
C:Genetics:  
A:Gene: HLA-B-4801  
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology  
C:Keywords: glycoprotein; transmembrane protein  
F:1-17/Domain: signal sequence #status predicted <SIG>  
F:17-354/Product: class I histocompatibility antigen HLA-B-4801 #status predicted <WA  
F:17-299/Domain: extracellular #status predicted <EXT>  
F:107-198/Domain: alpha-2 #status predicted <EX2>  
F:212-277/Domain: immunoglobulin homology <IMM>  
F:300-323/Domain: transmembrane #status predicted <TMN>  
F:324-354/Domain: intracellular #status predicted <INT>  
F:102/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:117-180,219-275/Disulfide bonds: #status predicted

Query Match 100.0%; Score 51; DB 2; Length 354;  
Best Local Similarity 100.0%; Pred. No. 0.012;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RESLNRLGY 10  
| | | | | | | | | |  
Db 91 RESLNRLGY 100

Search completed: February 7, 2000, 18:04:37  
Job time: 22203 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run On: February 8, 2000, 01:25:57 ; Search time 68.63 Seconds  
(without alignments)  
4.352 Million cell updates/sec

Title: US-08-653-294-24  
Perfect score: 51  
Sequence: 1 RESLNRLRGY 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 82229 seqs, 29864866 residues

Total number of hits satisfying chosen parameters: 82229

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database : SwissProt\_38.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	51	100.0	270	1	1B33_HUMAN
2	51	100.0	354	1	1B24_HUMAN
3	51	100.0	362	1	1A02_PANTR
4	51	100.0	362	1	1B02_HUMAN
5	51	100.0	362	1	1B02_PANTR
6	51	100.0	362	1	1B04_HUMAN
7	51	100.0	362	1	1B07_HUMAN
8	51	100.0	362	1	1B08_HUMAN
9	51	100.0	362	1	1B10_HUMAN
10	51	100.0	362	1	1B11_HUMAN
11	51	100.0	362	1	1B12_HUMAN
12	51	100.0	362	1	1B13_HUMAN
13	51	100.0	362	1	1B21_HUMAN
14	51	100.0	362	1	1B22_HUMAN
15	51	100.0	362	1	1B23_HUMAN
16	51	100.0	362	1	1B25_HUMAN
17	51	100.0	362	1	1B26_HUMAN
18	51	100.0	362	1	1B27_HUMAN
19	51	100.0	362	1	1B28_HUMAN
20	51	100.0	362	1	1B31_HUMAN
21	51	100.0	362	1	1B32_HUMAN
22	51	100.0	362	1	1B34_HUMAN
23	51	100.0	362	1	1B35_HUMAN
24	51	100.0	362	1	1B36_HUMAN
25	51	100.0	362	1	1B38_HUMAN
26	51	100.0	362	1	1B39_HUMAN
27	51	100.0	362	1	1B43_HUMAN
28	51	100.0	362	1	1B46_HUMAN
29	51	100.0	362	1	1B48_HUMAN
30	51	100.0	362	1	1B55_HUMAN
31	51	100.0	362	1	1B56_HUMAN
32	51	100.0	362	1	1B57_HUMAN
33	51	100.0	362	1	1B58_HUMAN
34	51	100.0	362	1	1B59_HUMAN

35	51	100.0	362	1	1B63_HUMAN	P30498	homo sapien
36	44	86.3	342	1	1CXX_HUMAN	P10321	homo sapien
37	44	86.3	362	1	1B44_HUMAN	P30484	homo sapien
38	44	86.3	366	1	1C01_HUMAN	P30499	homo sapien
39	44	86.3	366	1	1C01_PANTR	P30686	pan troglod
40	44	86.3	366	1	1C02_HUMAN	P30500	homo sapien
41	44	86.3	366	1	1C05_HUMAN	P04222	homo sapien
42	44	86.3	366	1	1C06_HUMAN	P30503	homo sapien
43	44	86.3	366	1	1C11_HUMAN	Q28631	homo sapien
44	44	86.3	366	1	1C12_HUMAN	P30505	homo sapien
45	44	86.3	366	1	1C13_HUMAN	P30506	homo sapien

## ALIGNMENTS

RESULT 1  
ID 1B33\_HUMAN STANDARD; PRT; 270 AA.  
AC P01890;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 01-APR-1993 (Rel. 25, Last annotation update)  
DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, BW-60(B-40) B\*4001 ALPHA CHAIN  
DE (FRAGMENT).  
GN HLA-B OR HLAB.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE.  
RX MEDLINE: 84000412.  
RA LOPEZ DE CASTRO J.A., BRAGADO R., STRONG D.M., STROMINGER J.L.;  
RT "Primary structure of papain-solubilized human histocompatibility  
antigen HLA-B\*40 (-Bw60). An outline of alloantigenic determinants";  
RL Biochemistry 22:3961-3969(1983).  
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO  
THE IMMUNE SYSTEM.  
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
MICROGLOBULIN).  
DR PIR; A02186; HLH40.  
DR HSSP; P30460; IAGB.  
DR MIM; 142830;  
DR PROSITE; PS00290; IG\_MHC; 1.  
DR PFAM; PF00047; Ig; 1.  
DR PFAM; PF00129; MHC\_I; 1.  
KW MHC I; Glycoprotein.  
FT DOMAIN 1 90  
FT DOMAIN 91 181  
FT DOMAIN 182 >270  
FT CARBOHYD 86 86  
FT DISULFID 101 163  
FT DISULFID 202 258  
FT NON\_TER 270 270  
SQ SEQUENCE 270 AA; BFE44EFF CRC32;  
Query Match 100.0%; Score 51; DB 1; Length 270;  
Best Local Similarity 100.0%; Pred. No. 0.0048;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RESLNRLRGY 10  
DB 75 RESLNRLRGY 84

RESULT 2  
ID 1B24\_HUMAN STANDARD; PRT; 354 AA.  
AC P30470;  
DT 01-APR-1993 (Rel. 25, Created)  
DT 01-APR-1993 (Rel. 25, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-35 B\*3504 ALPHA CHAIN  
DE PRECURSOR (FRAGMENT).

GN HLA-B OR HLAB.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

OC Eutheria; Primates; Catarrhini; Homidae; Homo.

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE; 92269956.

RA WATKINS D.I., MCADAM S.N., LIU X., STANG C.R., MILFORD E.L.,

RA LEVINE C.G., GABER T.L., DOON A.L., LORD C.I., GHIM S.H.,

RA TROUP G.M., HUGHES A.L., LETVIN N.L.;

"New recombinant HLA-B alleles in a tribe of South American

RT Amerindians indicate rapid evolution of MHC class I loci."

RL Nature 337:333(1992).

CC -1- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO

CC THE IMMUNE SYSTEM.

CC -1- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-

CC MICROGLOBULIN).

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC

CC EMBL; M86403; ; NOT\_ANNOTATED\_CDS.

DR HSP; P30685; IAIN.

DR MIM; 142830; ;

DR PROSITE; PS00290; IG\_MHC; 1.

DR PFAM; PF00047; Ig; 1.

DR PFAM; PF00129; MHC\_I; 1.

KW MHC I; Transmembrane; Glycoprotein; Signal.

FT NON\_TER 1 1

FT SIGNAL <1 16

FT CHAIN 17 354

FT HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,

FT B-35 B\*3504 ALPHA CHAIN.

FT EXTRACELLULAR ALPHA-1.

FT DOMAIN 17 106

FT DOMAIN 107 198

FT DOMAIN 199 290

FT DOMAIN 291 300

FT DOMAIN 301 324

FT TRANSMEM 301 324

FT DOMAIN 325 354

FT CYTOPLASMIC TAIL.

FT CARBOHYD 102 102

FT BY SIMILARITY.

FT DISULFID 117 180

FT BY SIMILARITY.

FT CARBOHYD 219 275

FT BY SIMILARITY.

FT SEQUENCE 354 AA; 39617 MW; 6564795A CRC32;

QY

1 RESLNRLRGY 10

|||||

Db 91 RESLNRLRGY 100

RESULT 3

1A02\_PANTR

ID 1A02\_PANTR STANDARD; PRT; 362 AA.

AC P16210;

DT 01-APR-1990 (Rel. 14, Created)

DT 01-APR-1990 (Rel. 14, Last sequence update)

DT 01-APR-1993 (Rel. 25, Last annotation update)

DE CHLA CLASS I HISTOCOMPATIBILITY ANTIGEN, A-5 ALPHA CHAIN PRECURSOR.

OS Pan troglodytes (Chimpanzee).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

OC Eutheria; Primates; Catarrhini; Homidae; Pan.

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE; 90201944.  
RA LAWLOR D.A., WARREN E., WARD F.E., PARHAM P.;  
RT "Comparison of class I MHC alleles in humans and apes."  
RL Immunol. Rev. 113:147-185(1990).  
CC -1- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO  
CC THE IMMUNE SYSTEM.  
CC -1- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
CC MICROGLOBULIN).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC

CC EMBL; M30679; AAA87971.1; ;

DR HSP; P30685; IAIN.

DR PROSITE; PS00290; IG\_MHC; 1.

DR PFAM; PF00047; Ig; 1.

DR PFAM; PF00129; MHC\_I; 1.

KW MHC I; Transmembrane; Glycoprotein; Signal.

FT SIGNAL 1 24

FT CHAIN 25 362

FT CHLA CLASS I HISTOCOMPATIBILITY ANTIGEN,

FT A-5 ALPHA CHAIN.

FT EXTRACELLULAR ALPHA-1.

FT DOMAIN 25 114

FT DOMAIN 115 206

FT DOMAIN 207 298

FT DOMAIN 299 308

FT DOMAIN 309 332

FT TRANSMEM 309 332

FT DOMAIN 333 362

FT CYTOPLASMIC TAIL.

FT DISULFID 125 188

FT BY SIMILARITY.

FT DISULFID 227 283

FT BY SIMILARITY.

FT CARBOHYD 110 110

FT SEQUENCE 362 AA; 40487 MW; 9756CE8A CRC32;

QY

1 RESLNRLRGY 10

|||||

Db 99 RESLNRLRGY 108

RESULT 4

1B02\_HUMAN

ID 1B02\_HUMAN STANDARD; PRT; 362 AA.

AC P01889;

DT 21-JUL-1986 (Rel. 01, Created)

DT 01-FEB-1991 (Rel. 17, Last sequence update)

DT 01-OCT-1996 (Rel. 34, Last annotation update)

DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-7 B\*0702 ALPHA CHAIN

DE PRECURSOR (B7.2).

GN HLA-B OR HLAB.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

OC Eutheria; Primates; Catarrhini; Homidae; Homo.

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE; 90207291.

RA ENNIS P.D., ZEMMOUR J., SALTER R.D., PARHAM P.;

"Rapid cloning of HLA-A,B CDNA by using the polymerase chain

RT reaction: frequency and nature of errors produced in amplification."

RL Proc. Natl. Acad. Sci. U.S.A. 87:2833-2837(1990).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE; 90315860.

RA PARHAM P., BENJAMIN R.J., CHEN B.P., CLAYBERGER C., ENNIS P.D.,

RA KRENSKY A.M., LAWLOR D.A., LITTMAN D.R., NORMENT A.M., ORR H.T.,

RA SALTER R.D., ZEMMOUR J.

RT "Diversity of class I HLA molecules: functional and evolutionary  
 RL interactions with T cells.";  
 RN Cold Spring Harb. Symp. Quant. Biol. 54:529-543(1989).  
 RX [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 85287366.  
 RA SOOD A.K., PAN J., BIRO P.A., PEREIRA D., SRIVASTAVA R., REDDY V.B.,  
 RA DUCMAN B.W., WEISSMAN S.M.;  
 RT "Structure and polymorphism of class I MHC antigen mRNA.";  
 RL Immunogenetics 22:101-121(1985).  
 [4]  
 RP SEQUENCE FROM N.A.  
 RA ELLERSON M.E., ZHANG L., HILDEBRAND W.H.;  
 RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.  
 [5]  
 RP SEQUENCE OF 25-295.  
 RX MEDLINE: 80088278.  
 RA ORR H.T., LOPEZ DE CASTRO J.A., LANCET D., STROMINGER J.L.;  
 RT "Complete amino acid sequence of a papain-solubilized human  
 RT histocompatibility antigen, HLA-B\*7.2. Sequence determination and  
 RT search for homologies.";  
 RL Biochemistry 18:5711-5720(1979).  
 CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO  
 CC THE IMMUNE SYSTEM.  
 CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
 CC MICROGLOBULIN).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL: M32317; AAA36230.1; -;  
 DR EMBL: M16102; AAA59622.1; AUT\_SEQ.  
 DR EMBL: U29057; AAA91229.1; -;  
 DR PIR: A02185; HLHUB7.  
 DR PIR: B35997; B35997.  
 DR HSP: P30460; LAGB.  
 DR MIM: 142830; -;  
 DR PROSITE: PS00290; IG\_MHC; 1.  
 DR PFAM: PF00047; ig; 1.  
 DR PFAM: PF01129; MHC\_I; 1.  
 KW MHC I; Transmembrane; Glycoprotein; Signal.  
 FT SIGNAL 1 24  
 FT CHAIN 25 362  
 FT HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,  
 FT B\*7 B\*0702 ALPHA CHAIN.  
 FT DOMAIN 25 114  
 FT DOMAIN 115 206  
 FT DOMAIN 207 298  
 FT DOMAIN 299 309  
 FT DOMAIN 309 333  
 FT TRANSMEM 333 362  
 FT DOMAIN 334 362  
 FT CARBOHYD 110 110  
 FT DISULFID 125 188  
 FT DISULFID 227 283  
 FT CONFLICT 15 18  
 FT CONFLICT 266 266  
 FT CONFLICT 268 268  
 FT CONFLICT 297 297  
 FT CONFLICT 314 315  
 FT CONFLICT 314 315  
 FT SEQUENCE 362 AA; 40460 MW; 8782ED84 CRC32;  
 AALA -> GPW (IN REF. 3).  
 Q -> E (IN REF. 5).  
 W -> S (IN REF. 3).  
 R -> G (IN REF. 3).  
 GL -> RP (IN REF. 3).  
 Query Match 100.0%; Score 51; DB 1; Length 362;  
 Best Local Similarity 100.0%; Pred. No. 0.0066;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RESLNRNGY 10  
 DB 99 RESLNRNGY 108  
 RESULT 6  
 1B04\_HUMAN STANDARD; PRT; 362 AA.  
 AC P30460;

RESULT 5  
 1B02\_PANTR STANDARD; PRT; 362 AA.  
 ID 1B02\_PANTR  
 AC P13751;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 01-APR-1993 (Rel. 25, Last annotation update)  
 DE CHLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-2 ALPHA CHAIN PRECURSOR.  
 OS Pan troglodytes (Chimpanzee).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominiidae; Pan.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 89030641.  
 RA MAYER W.E., JONKER M., KLEIN D., IVANYI P., VAN SEVENTER G.,  
 RA KLEIN J.;  
 RT "Nucleotide sequences of chimpanzee MHC class I alleles: evidence for  
 RT trans-species mode of evolution.";  
 RL EMBO J. 7:2765-2774(1988).  
 RN [2]  
 RP REVISIONS.  
 RA MAYER W.;  
 RL Submitted (FEB-1989) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO  
 CC THE IMMUNE SYSTEM.  
 CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
 CC MICROGLOBULIN).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL: X13116; CAA31508.1; -;  
 DR PIR: S03538; S03538.  
 DR HSP: P30685; IAIN.  
 DR PROSITE: PS00290; IG\_MHC; 1.  
 DR PFAM: PF00047; ig; 1.  
 DR PFAM: PF01129; MHC\_I; 1.  
 KW MHC I; Transmembrane; Glycoprotein; Signal.  
 FT SIGNAL 1 24  
 FT CHAIN 25 362  
 FT CHLA CLASS I HISTOCOMPATIBILITY ANTIGEN,  
 FT B-2 ALPHA CHAIN.  
 FT DOMAIN 25 114  
 FT DOMAIN 115 206  
 FT DOMAIN 207 298  
 FT DOMAIN 299 308  
 FT TRANSMEM 309 332  
 FT DOMAIN 333 362  
 FT DISULFID 125 188  
 FT DISULFID 227 283  
 FT CARBOHYD 110 110  
 FT SEQUENCE 362 AA; 40488 MW; 4BF65A6C CRC32;  
 Query Match 100.0%; Score 51; DB 1; Length 362;  
 Best Local Similarity 100.0%; Pred. No. 0.0066;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RESLNRNGY 10  
 DB 99 RESLNRNGY 108  
 RESULT 6  
 1B04\_HUMAN STANDARD; PRT; 362 AA.  
 ID 1B04\_HUMAN  
 AC P30460;

```
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-8 B*0801 ALPHA CHAIN
DE PRECURSOR.
GN HLA-B OR HLAB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 89235215.
RA PARHAM P., LAWLER D.A., LOMEN C.E., ENNIS P.D.;
RT "Diversity and diversification of HLA-A,B,C alleles.";
RL J. Immunol. 142:3937-3950(1989).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 25-300.
RX MEDLINE; 97130420.
RA REID S.W., MCADAM S., SMITH K.J., KLENERMAN P., O'CALLAGHAN C.A.,
RA HARLOS K., JAKOBSEN B.K., MCMICHAEL A.J., BELL J.I., STUART D.I.,
RA JONES E.Y.;
RT "Antagonist HIV-1 Gag peptides induce structural changes in HLA B8.";
RL J. Exp. Med. 184:2279-2286(1996).
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
THE IMMUNE SYSTEM.
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
MICROGLOBULIN).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; M24036; AAA52662.1; -.
DR PDB; IAGB; 16-JUN-97.
DR PDB; IAGC; 16-JUN-97.
DR PDB; IAGD; 16-JUN-97.
DR PDB; IAGE; 16-JUN-97.
DR PDB; IAGF; 16-JUN-97.
DR MIM; 142830; -.
DR PROSITE; PS00290; IG_MHC; 1.
DR PFAM; PF00047; Ig; 1.
DR PFAM; PF00129; MHC_I; 1.
DR MHC I; Transmembrane; Glycoprotein; Signal; 3D-structure.
FT SIGNAL 1 24
FT CHAIN 25 362
FT DOMAIN 25 114 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
FT DOMAIN 115 206 B-8 B*0801 ALPHA CHAIN.
FT DOMAIN 207 298 EXTRACELLULAR ALPHA-1.
FT DOMAIN 299 309 EXTRACELLULAR ALPHA-2.
FT TRANSMEM 310 333 EXTRACELLULAR ALPHA-3.
FT DOMAIN 334 362 CONNECTING PEPTIDE.
FT CARBOHYD 110 110 CYTOPLASMIC TAIL.
FT DISULFID 125 188 BY SIMILARITY.
FT DISULFID 227 283 BY SIMILARITY.
SQ SEQUENCE 362 AA; 40331 MW; 1467B8EB CRC32;

Query Match 100.0%; Score 51; DB 1; Length 362;
Best Local Similarity 100.0%; Pred. No. 0.0066;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RESLNRLRGY 10
DB 99 RESLNRLRGY 108

RESULT 7
1B07_HUMAN
ID 1B08_HUMAN STANDARD; PRT; 362 AA.
AC P30463;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-14 B*1401 ALPHA CHAIN
DE PRECURSOR.
GN HLA-B OR HLAB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 89235215.
RA PARHAM P., LAWLER D.A., LOMEN C.E., ENNIS P.D.;
RT "Diversity and diversification of HLA-A,B,C alleles.";
RL J. Immunol. 142:3937-3950(1989).
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
THE IMMUNE SYSTEM.
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
MICROGLOBULIN).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; M24040; AAA59661.1; -.
DR HSP; P30460; IAGB.
DR MIM; 142830; -.
DR PROSITE; PS00290; IG_MHC; 1.
DR PFAM; PF00047; Ig; 1.
DR PFAM; PF00129; MHC_I; 1.
DR MHC I; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 24
FT CHAIN 25 362 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
FT DOMAIN 25 114 B-14 B*1401 ALPHA CHAIN.
FT DOMAIN 115 206 EXTRACELLULAR ALPHA-1.
FT DOMAIN 207 298 EXTRACELLULAR ALPHA-2.
FT DOMAIN 299 309 EXTRACELLULAR ALPHA-3.
FT TRANSMEM 310 333 CONNECTING PEPTIDE.
FT DOMAIN 334 362 CYTOPLASMIC TAIL.
FT CARBOHYD 110 110 BY SIMILARITY.
FT DISULFID 125 188 BY SIMILARITY.
FT DISULFID 227 283 BY SIMILARITY.
SQ SEQUENCE 362 AA; 40358 MW; 9BED8199 CRC32;

Query Match 100.0%; Score 51; DB 1; Length 362;
Best Local Similarity 100.0%; Pred. No. 0.0066;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RESLNRLRGY 10
DB 99 RESLNRLRGY 108

RESULT 8
1B08_HUMAN
ID 1B08_HUMAN STANDARD; PRT; 362 AA.
AC P30463;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-65(B-14) B*1402 ALPHA CHAIN
DE PRECURSOR.
GN HLA-B OR HLAB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
```



OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 (1)  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE; 89235215.  
 RX PARHAM P., LAWLER D.A., LOMEN C.E., ENNIS P.D.;  
 RA "Diversity and diversification of HLA-A,B,C alleles."  
 RT J. Immunol. 142:3937-3950(1989).  
 RL  
 CC -1- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO  
 CC THE IMMUNE SYSTEM.  
 CC -1- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
 CC MICROGLOBULIN).  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; M24032; AAA59664.1; -;  
 DR HSP; P30460; IAGB.  
 DR MIN; 142830; -;  
 DR PROSITE; PS00290; IG\_MHC; 1.  
 DR PFAM; PF00047; Ig; 1.  
 DR PFAM; PF00129; MHC\_I; 1.  
 KW MHC I; Transmembrane; Glycoprotein; Signal.  
 FT SIGNAL 1 24  
 FT CHAIN 25 362 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,  
 FT BW-65(B-14) B\*1402 ALPHA CHAIN.  
 FT DOMAIN 25 114 EXTRACELLULAR ALPHA-1.  
 FT DOMAIN 115 206 EXTRACELLULAR ALPHA-2.  
 FT DOMAIN 207 298 EXTRACELLULAR ALPHA-3.  
 FT DOMAIN 299 309 CONNECTING PEPTIDE.  
 FT TRANSMEM 310 333  
 FT DOMAIN 334 362 CYTOPLASMIC TAIL.  
 FT CARBOHYD 110 110 BY SIMILARITY.  
 FT DISULFID 125 188 BY SIMILARITY.  
 FT DISULFID 227 283 BY SIMILARITY.  
 SQ SEQUENCE 362 AA; 40342 MW; B568AC9E CRC32;  
 Query Match 100.0%; Score 51; DB 1; Length 362;  
 Best Local Similarity 100.0%; Pred. No. 0.0066;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RESLRNLRGY 10  
 Db 99 RESLRNLRGY 108  
 RESULT 9  
 LB10\_HUMAN  
 ID 1B10\_HUMAN STANDARD; PRT; 362 AA.  
 AC P30464;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, BW-75(B-15) B\*1502 ALPHA CHAIN  
 DE PRECURSOR.  
 GN HLA-B OR HLAB.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 92196792.  
 RA "The HLA-B\*75 subtype of B15: molecular characterization and  
 RT comparison with crossreacting antigens."  
 RL Tissue Antigens 38:186-190(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.

RC TISSUE-BLOOD;  
 RX MEDLINE; 96369309.  
 RA LIN L., TOKUNAGA K., TANAKA H., NAKAJIMA F., IMANISHI T.,  
 RA KASHIWASE K., BANNAI M., MIZUNO S., AKAZA T., TADOKORO K.,  
 RA SHIBATA Y., JUJI T.;  
 RT "Further molecular diversity in the HLA-B\*15 group."  
 RL Tissue Antigens 47:265-274(1996).  
 CC -1- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO  
 CC THE IMMUNE SYSTEM.  
 CC -1- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
 CC MICROGLOBULIN).  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; M75138; AAA59630.1; -;  
 DR EMBL; D50293; BAA08824.1; -;  
 DR HSP; P30685; IAIN.  
 DR MIN; 142830; -;  
 DR PROSITE; PS00290; IG\_MHC; 1.  
 DR PFAM; PF00047; Ig; 1.  
 DR PFAM; PF00129; MHC\_I; 1.  
 KW MHC I; Transmembrane; Glycoprotein; Signal.  
 FT SIGNAL 1 24  
 FT CHAIN 25 362 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,  
 FT BW-75(B-15) B\*1502 ALPHA CHAIN.  
 FT DOMAIN 25 114 EXTRACELLULAR ALPHA-1.  
 FT DOMAIN 115 206 EXTRACELLULAR ALPHA-2.  
 FT DOMAIN 207 298 EXTRACELLULAR ALPHA-3.  
 FT DOMAIN 299 309 CONNECTING PEPTIDE.  
 FT TRANSMEM 310 333  
 FT DOMAIN 334 362 CYTOPLASMIC TAIL.  
 FT CARBOHYD 110 110 BY SIMILARITY.  
 FT DISULFID 125 188 BY SIMILARITY.  
 FT DISULFID 227 283 BY SIMILARITY.  
 SQ SEQUENCE 362 AA; 40338 MW; 8CF9BCD0 CRC32;  
 Query Match 100.0%; Score 51; DB 1; Length 362;  
 Best Local Similarity 100.0%; Pred. No. 0.0066;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RESLRNLRGY 10  
 Db 99 RESLRNLRGY 108  
 RESULT 10  
 IB11\_HUMAN  
 ID 1B11\_HUMAN STANDARD; PRT; 362 AA.  
 AC P30465;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, BW-72(BW-70) B\*1503 ALPHA  
 DE CHAIN PRECURSOR.  
 GN HLA-B OR HLAB.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 93056508.  
 RA MADRIGAL J.A., BELICH M.P., HILDEBRAND W.H., BENJAMIN R.J.,  
 RA LITTLE A.-M., ZEMMOUR J., ENNIS P.D., WARD F.E., PETZL-ERLER M.L.,  
 RA MARTELL R.W., DU TOIT E.D., PARHAM P.;  
 RT "Distinctive HLA-A,B antigens of black populations formed by  
 RT interallelic conversion."

RL J. Immunol. 149:3411-3415(1992).  
 CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO  
 CC THE IMMUNE SYSTEM.  
 CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
 CC MICROGLOBULIN).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: X61709; CAA43878.1; -  
 DR PIR: S16789; S16789.  
 DR HSP: P30685; IAIN.  
 DR MIM: 142830; -  
 DR PROSITE: PS00290; IG\_MHC; 1.  
 DR PFAM: PF00047; ig; 1.  
 DR PFAM: PF00129; MHC\_I; 1.  
 KW MHC I; Transmembrane; Glycoprotein; Signal.  
 FT SIGNAL 1 24  
 FT CHAIN 25 362 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,  
 FT BW-72(BW-70) B\*1503 ALPHA CHAIN.  
 FT DOMAIN 25 114 EXTRACELLULAR ALPHA-1.  
 FT DOMAIN 115 206 EXTRACELLULAR ALPHA-2.  
 FT DOMAIN 207 298 EXTRACELLULAR ALPHA-3.  
 FT DOMAIN 299 309 CONNECTING PEPTIDE.  
 FT TRANSMEM 310 333  
 FT DOMAIN 334 362 CYTOPLASMIC TAIL.  
 FT CARBOHYD 110 110 BY SIMILARITY.  
 FT DISULFID 125 188 BY SIMILARITY.  
 FT DISULFID 227 283 BY SIMILARITY.  
 SQ SEQUENCE 362 AA: 40387 MW: 99D70546 CRC32;

Query Match 100.0%; Score 51; DB 1; Length 362;  
 Best Local Similarity 100.0%; Pred. No. 0.0066;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RESLRNLRGY 10  
 | | | | | | | | | |  
 Db 99 RESLRNLRGY 108

RESULT 11  
 ID 1B12\_HUMAN STANDARD; PRT; 362 AA.  
 AC P30513;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-62 B\*1504 ALPHA CHAIN  
 DE PRECURSOR.  
 GN HLA-B OR HLAB.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 [1]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE; 92269956.  
 RA WATKINS D.I., MCADAM S.N., LIU X., STANG C.R., MILFORD E.L.,  
 RA LEVINE C.G., GARBER T.L., DOGON A.L., LORD C.I., GHIM S.H.,  
 RA TROUP G.M., HUGHES A.L., LEVIN N.I.;  
 RT "New recombinant HLA-B alleles in a tribe of South American  
 RT Amerindians indicate rapid evolution of MHC class I loci."  
 RL Nature 357:329-333(1992).  
 [2]  
 RN SEQUENCE FROM N.A.  
 RP TISSUE-BLOOD:  
 RC RAMOS M., BARBER D.F., LAYRISSE Z., DE CASTRO J.A.;  
 RL Submitted (Oct-1996) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO  
 CC THE IMMUNE SYSTEM.  
 CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
 CC MICROGLOBULIN).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: M84382; AAA59632.1; -  
 DR EMBL: U70528; AAB16918.1; -  
 DR PIR: S24433; S24433.  
 DR HSP: P30685; IAIN.  
 DR MIM: 142830; -  
 DR PROSITE: PS00290; IG\_MHC; 1.  
 DR PFAM: PF00047; ig; 1.  
 DR PFAM: PF00129; MHC\_I; 1.  
 KW MHC I; Transmembrane; Glycoprotein; Signal.  
 FT SIGNAL 1 24  
 FT CHAIN 25 362 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,  
 FT B-62 B\*1504 ALPHA CHAIN.  
 FT DOMAIN 25 114 EXTRACELLULAR ALPHA-1.  
 FT DOMAIN 115 206 EXTRACELLULAR ALPHA-2.  
 FT DOMAIN 207 298 EXTRACELLULAR ALPHA-3.  
 FT DOMAIN 299 309 CONNECTING PEPTIDE.  
 FT TRANSMEM 310 333  
 FT DOMAIN 334 362 CYTOPLASMIC TAIL.  
 FT CARBOHYD 110 110 BY SIMILARITY.  
 FT DISULFID 125 188 BY SIMILARITY.  
 FT DISULFID 227 283 BY SIMILARITY.  
 SQ SEQUENCE 362 AA: 40406 MW: E80FC24C CRC32;

Query Match 100.0%; Score 51; DB 1; Length 362;  
 Best Local Similarity 100.0%; Pred. No. 0.0066;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RESLRNLRGY 10  
 | | | | | | | | | |  
 Db 99 RESLRNLRGY 108

RESULT 12  
 ID 1B13\_HUMAN STANDARD; PRT; 362 AA.  
 AC P30466;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-18 B\*1801 ALPHA CHAIN  
 DE PRECURSOR.  
 GN HLA-B OR HLAB.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 [1]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE; 89235215.  
 RA PARHAM P., LAWLER D.A., LOMEN C.E., ENNIS P.D.;  
 RT "Diversity and diversification of HLA-A,B,C alleles."  
 RL J. Immunol. 142:3937-3950(1989).  
 CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO  
 CC THE IMMUNE SYSTEM.  
 CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
 CC MICROGLOBULIN).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; M24039; AAA59662.1; -  
DR HSSP; P30460; 1AGB.  
DR MIN; 142830; -  
DR PROSITE; PS00290; IG\_MHC; 1.  
DR PFAM; PF00047; Ig; 1.  
DR PFAM; PF00129; MHC\_I; 1.  
DR PFAM; PF00129; MHC\_I; 1.  
KW MHC I; Transmembrane; Glycoprotein; Signal.  
FT SIGNAL 1 24  
FT CHAIN 25 362  
FT  
FT HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,  
FT B-18 B\*1801 ALPHA CHAIN.  
FT  
FT DOMAIN 25 114  
FT DOMAIN 115 206  
FT DOMAIN 207 298  
FT DOMAIN 299 309  
FT TRANSMEM 310 333  
FT DOMAIN 334 362  
FT CARBOHYD 110 110  
FT DISULFID 125 188  
FT DISULFID 227 283  
SQ SEQUENCE 362 AA; 40275 MW; 038EC3FC CRC32;

Query Match 100.0%; Score 51; DB 1; Length 362;  
Best Local Similarity 100.0%; Pred. No. 0.0066;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RESRLNLRGY 10  
| | | | | | | |  
Db 99 RESRLNLRGY 108

RESULT 13  
1B21\_HUMAN  
ID 1B21\_HUMAN STANDARD; PRT; 362 AA.  
AC P30685;  
DT 01-APR-1993 (Rel. 25, Created)  
DT 01-APR-1993 (Rel. 25, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-35 B\*3501 ALPHA CHAIN  
DE PRECURSOR.  
GN HLA-B OR HLAB.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
[1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 89339610.  
RA Ooba T., HAYASHI H., KARAKI S., TANABE M., KANO K., TAKIGUCHI M.;  
RT "The structure of HLA-B\*35 suggests that it is derived from HLA-B\*58  
RT by two genetic mechanisms.";  
RL Immunogenetics 30:76-80(1989).  
[2]  
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 25-300.  
RX MEDLINE; 96209671.  
RA SMITH K.J., REID S.W., STUART D.I., MCMICHAEL A.J., JONES E.Y.,  
RA BELL J.I.;  
RT "An altered position of the alpha 2 helix of MHC class I is revealed  
RT by the crystal structure of HLA-B\*3501.";  
RL Immunity 4:203-214(1996).  
[3]  
RP X-RAY CRYSTALLOGRAPHY (3.2 ANGSTROMS).  
RA MENSSSEN R., ORTH P., ZIEGLER A., SAENGER W.;  
RL Submitted (APR-1998) to the PDB data bank.  
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO  
CC THE IMMUNE SYSTEM.  
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
CC MICROGLOBULIN).  
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; M28115; AAA59617.1; -  
DR EMBL; M28109; AAA59617.1; JOINED.  
DR EMBL; M28110; AAA59617.1; JOINED.  
DR EMBL; M28111; AAA59617.1; JOINED.  
DR EMBL; M28112; AAA59617.1; JOINED.  
DR EMBL; M28113; AAA59617.1; JOINED.  
DR EMBL; M28114; AAA59617.1; JOINED.  
DR PIR; A45880; A45880.  
DR PDB; 1A9B; 08-APR-98.  
DR PDB; 1A9B; 18-NOV-98.  
DR PDB; 1A9E; 18-NOV-98.  
DR MIN; 142830; -  
DR PROSITE; PS00290; IG\_MHC; 1.  
DR PFAM; PF00047; Ig; 1.  
DR PFAM; PF00129; MHC\_I; 1.  
KW MHC I; Transmembrane; Glycoprotein; Signal; 3D-structure.  
FT SIGNAL 1 24  
FT CHAIN 25 362  
FT  
FT DOMAIN 25 114  
FT DOMAIN 115 206  
FT DOMAIN 207 298  
FT DOMAIN 299 308  
FT TRANSMEM 309 332  
FT DOMAIN 333 362  
FT CARBOHYD 110 110  
FT DISULFID 125 188  
FT DISULFID 227 283  
SQ SEQUENCE 362 AA; 40455 MW; AEC1C675 CRC32;

Query Match 100.0%; Score 51; DB 1; Length 362;  
Best Local Similarity 100.0%; Pred. No. 0.0066;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RESRLNLRGY 10  
| | | | | | | |  
Db 99 RESRLNLRGY 108

RESULT 14  
1B22\_HUMAN  
ID 1B22\_HUMAN STANDARD; PRT; 362 AA.  
AC P30468;  
DT 01-APR-1993 (Rel. 25, Created)  
DT 01-APR-1993 (Rel. 25, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-35 B\*3502 ALPHA CHAIN  
DE PRECURSOR.  
GN HLA-B OR HLAB.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
[1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 91365651.  
RA CHERTKOFF L.P., HERRERA M., FAIBOIM L., SATZ M.L.;  
RT "Complete nucleotide sequence of a genomic clone encoding HLA-B\*35  
RT isolated from a Caucasian individual of Hispanic origin.  
RT Identification of a new variant of HLA-B\*35.";  
RL Hum. Immunol. 31:153-158(1991).  
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO  
CC THE IMMUNE SYSTEM.  
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
CC MICROGLOBULIN).  
CC -----

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 -----

DR EMBL; M63454; AAS9682.1; -  
 DR HSP; P30685; IAIN.  
 DR MIM; 142830; -  
 DR PROSITE; PS00290; IG\_MHC; 1.  
 DR PFAM; PF00047; ig; 1.  
 DR PFAM; PF00129; MHC\_I; 1.  
 KW MHC I; Transmembrane; Glycoprotein; Signal.  
 FT SIGNAL 1 24  
 FT CHAIN 25 362  
 FT DOMAIN 25 114  
 FT DOMAIN 115 206  
 FT DOMAIN 207 298  
 FT DOMAIN 299 308  
 FT TRANSMEM 309 332  
 FT DOMAIN 333 362  
 FT CARBOHYD 110 110  
 FT DISULFID 125 188  
 FT DISULFID 227 283  
 SQ SEQUENCE 362 AA; D7B5C2C9 CRC32;

Query Match 100.0%; Score 51; DB 1; Length 362;

Best Local Similarity 100.0%; Pred. No. 0.0066;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RESLRNLRGY 10

DB 99 RESLRNLRGY 108

# RESULT 15

1B23\_HUMAN STANDARD; PRT; 362 AA.  
 ID 1B23\_HUMAN  
 AC P30469;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-35 B\*3503 ALPHA CHAIN  
 DE PRECURSOR.  
 GN HLA-B OR HLAB.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 [1]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE; 92176651.  
 RA ZEMMOUR J., LITTLE A.M., SCHENDEL D.J., PARHAM P.;  
 RT "The HLA-A,B 'negative' mutant cell line ClR expresses a novel  
 RT HLA-B35 allele, which also has a point mutation in the translation  
 RT initiation codon."  
 RL J. Immunol. 148:1941-1948(1992).  
 RN [2]  
 RN SEQUENCE FROM N.A.  
 RC TISSUE-BLOOD;  
 RX MEDLINE; 95279930.  
 RA BECK Y., SATZ L., TAKAMIYA Y., NAKAYAMA S., LING L., ISHIIKAWA Y.,  
 RA NAGAO T., UCHIDA H., TOKUNAGA K., MULLER C., JUJI T., TAKIGUCHI M.;  
 RT "Polymorphism of human minor histocompatibility antigens: T cell  
 RT recognition of human minor histocompatibility peptides presented by  
 RT HLA-B35 subtype molecules."  
 RL J. Exp. Med. 181:2037-2048(1995).  
 CC -1- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO  
 CC THE IMMUNE SYSTEM.

CC -1- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
 CC MICROGLOBULIN).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 -----

DR EMBL; M81798; AAS9684.1; -  
 DR EMBL; D50299; BAA0828.1; -  
 DR HSP; P30685; IAIN.  
 DR MIM; 142830; -  
 DR PROSITE; PS00290; IG\_MHC; 1.  
 DR PFAM; PF00047; ig; 1.  
 DR PFAM; PF00129; MHC\_I; 1.  
 KW MHC I; Transmembrane; Glycoprotein; Signal.  
 FT SIGNAL 1 24  
 FT CHAIN 25 362  
 FT DOMAIN 25 114  
 FT DOMAIN 115 206  
 FT DOMAIN 207 298  
 FT DOMAIN 299 308  
 FT TRANSMEM 309 332  
 FT DOMAIN 333 362  
 FT CARBOHYD 110 110  
 FT DISULFID 125 188  
 FT DISULFID 227 283  
 SQ SEQUENCE 362 AA; 40515 MW; 4222D30A CRC32;

Query Match 100.0%; Score 51; DB 1; Length 362;

Best Local Similarity 100.0%; Pred. No. 0.0066;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RESLRNLRGY 10

DB 99 RESLRNLRGY 108

Search completed: February 8, 2000, 01:25:58

Job time: 1558 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2000, 19:16:13 ; Search time 176.54 Seconds  
(without alignments)  
3.927 Million cell updates/sec

Title: US-08-653-294-24  
Perfect score: 51  
Sequence: 1 RESLNRLRG 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 225878 seqs, 69334122 residues

Total number of hits satisfying chosen parameters: 225878

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database : SPTREMBL\_12.\*  
1: sp.archaea.\*  
2: sp.bacteria.\*  
3: sp.fungi.\*  
4: sp.human.\*  
5: sp.invertebrate.\*  
6: sp.mammal.\*  
7: sp.mhc.\*  
8: sp.organelle.\*  
9: sp.phage.\*  
10: sp.plant.\*  
11: sp.rodent.\*  
12: sp.virus.\*  
13: sp.vertebrate.\*  
14: sp.unclassified.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result NO.	Score	Query Match	Length	DB ID	Description
1	51	100.0	81	7	O19523
2	51	100.0	81	7	O19525
3	51	100.0	81	7	O19527
4	51	100.0	81	7	O19529
5	51	100.0	81	7	O19531
6	51	100.0	81	7	O19533
7	51	100.0	81	7	O19535
8	51	100.0	81	7	O19537
9	51	100.0	83	7	O19539
10	51	100.0	83	7	O19541
11	51	100.0	83	7	O19543
12	51	100.0	89	4	Q9Y452
13	51	100.0	89	7	P79615
14	51	100.0	89	7	P79620
15	51	100.0	89	7	O19548
16	51	100.0	89	7	O19549
17	51	100.0	89	7	O19550
18	51	100.0	89	7	O19551
19	51	100.0	89	7	O19640
20	51	100.0	89	7	O77967

21	51	100.0	89	7	O78174	078174	homo sapien
22	51	100.0	89	7	O19638	O19638	homo sapien
23	51	100.0	89	7	P79487	P79487	homo sapien
24	51	100.0	89	7	O19567	O19567	homo sapien
25	51	100.0	89	7	O19576	O19576	homo sapien
26	51	100.0	89	7	O19577	O19577	homo sapien
27	51	100.0	89	7	O19581	O19581	homo sapien
28	51	100.0	89	7	O19583	O19583	homo sapien
29	51	100.0	112	7	O92671	O92671	homo sapien
30	51	100.0	134	7	O29677	O29677	homo sapien
31	51	100.0	137	7	O29660	O29660	homo sapien
32	51	100.0	137	7	O29663	O29663	homo sapien
33	51	100.0	137	7	O29665	O29665	homo sapien
34	51	100.0	137	7	O29664	O29664	homo sapien
35	51	100.0	147	7	O29707	O29707	homo sapien
36	51	100.0	165	7	P79665	P79665	homo sapien
37	51	100.0	165	7	P79666	P79666	homo sapien
38	51	100.0	166	7	O19777	O19777	homo sapien
39	51	100.0	166	7	O19778	O19778	homo sapien
40	51	100.0	172	7	O19776	O19776	homo sapien
41	51	100.0	176	7	O29662	O29662	homo sapien
42	51	100.0	179	7	O29676	O29676	homo sapien
43	51	100.0	180	7	O29678	O29678	homo sapien
44	51	100.0	180	7	O77960	O77960	homo sapien
45	51	100.0	181	7	O29716	O29716	homo sapien

## ALIGNMENTS

RESULT 1  
O19523 PRELIMINARY; PRT; 81 AA.  
AC O19523;  
DT 01-JAN-1998 (TRENBLrel. 05, Created)  
DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)  
DT 01-NOV-1998 (TRENBLrel. 08, Last annotation update)  
DE MHC CLASS I ANTIGEN HLA-B (FRAGMENT).  
GN HLA-B.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA CHANDANAYINGYONG D., SIRIKONG M., LUANGTRAKOOL K., SRINAK D.,  
RA RUNGRONG E., BECHANDRA S.;  
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF014769; AAB67807.1; -  
DR PFAM; PF00129; MHC\_I; 1.  
KW MHC.  
FT NON\_TER 1 1  
FT NON\_TER 81 81  
SQ SEQUENCE 81 AA; 9405 MW; 073087CE CRC32;  
Query Match 100.0%; Score 51; DB 7; Length 81;  
Best Local Similarity 100.0%; Pred. No. 0.013;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 RESLNRLRG 10  
Db 66 RESLNRLRG 75  
RESULT 2  
O19525 PRELIMINARY; PRT; 81 AA.  
AC O19525;  
DT 01-JAN-1998 (TRENBLrel. 05, Created)  
DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)  
DT 01-NOV-1998 (TRENBLrel. 08, Last annotation update)  
DE MHC CLASS I ANTIGEN HLA-B (FRAGMENT).  
GN HLA-B.

OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA CHANDANAYINGYONG D., SIRIKONG M., LUANGTRAKOOL K., SRINAK D.,  
 RA RUNGRONG E., BEJCHANDRA S.;  
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF014771; AAB67809.1; -  
 DR PFAM; PF00129; MHC\_I; 1.  
 KW MHC.  
 FT NON\_TER 1 1  
 FT NON\_TER 81 81  
 SQ SEQUENCE 81 AA; 9405 MW; 073087CE CRC32;

Query Match 100.0%; Score 51; DB 7; Length 81;  
 Best Local Similarity 100.0%; Pred. No. 0.013;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RESLRNLRGY 10  
 DB 66 RESLRNLRGY 75

RESULT 3  
 ID O19527 PRELIMINARY; PRT; 81 AA.  
 AC O19527;  
 DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
 DE MHC CLASS I ANTIGEN HLA-B (FRAGMENT).  
 GN HLA-B.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]

RP SEQUENCE FROM N.A.  
 RA CHANDANAYINGYONG D., SIRIKONG M., LUANGTRAKOOL K., SRINAK D.,  
 RA RUNGRONG E., BEJCHANDRA S.;  
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF014773; AAB67811.1; -  
 DR PFAM; PF00129; MHC\_I; 1.  
 KW MHC.  
 FT NON\_TER 1 1  
 FT NON\_TER 81 81  
 SQ SEQUENCE 81 AA; 9405 MW; 073087CE CRC32;

Query Match 100.0%; Score 51; DB 7; Length 81;  
 Best Local Similarity 100.0%; Pred. No. 0.013;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RESLRNLRGY 10  
 DB 66 RESLRNLRGY 75

RESULT 4  
 ID O19529 PRELIMINARY; PRT; 81 AA.  
 AC O19529;  
 DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
 DE MHC CLASS I ANTIGEN HLA-B (FRAGMENT).  
 GN HLA-B.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.

RA CHANDANAYINGYONG D., SIRIKONG M., LUANGTRAKOOL K., SRINAK D.,  
 RA RUNGRONG E., BEJCHANDRA S.;  
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF014775; AAB67813.1; -  
 DR PFAM; PF00129; MHC\_I; 1.  
 KW MHC.  
 FT NON\_TER 1 1  
 FT NON\_TER 81 81  
 SQ SEQUENCE 81 AA; 9405 MW; 073087CE CRC32;

Query Match 100.0%; Score 51; DB 7; Length 81;  
 Best Local Similarity 100.0%; Pred. No. 0.013;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RESLRNLRGY 10  
 DB 66 RESLRNLRGY 75

RESULT 5  
 ID O19531 PRELIMINARY; PRT; 81 AA.  
 AC O19531;  
 DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
 DE MHC CLASS I ANTIGEN HLA-B (FRAGMENT).  
 GN HLA-B.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]

RP SEQUENCE FROM N.A.  
 RA CHANDANAYINGYONG D., SIRIKONG M., LUANGTRAKOOL K., SRINAK D.,  
 RA RUNGRONG E., BEJCHANDRA S.;  
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF014777; AAB67815.1; -  
 DR PFAM; PF00129; MHC\_I; 1.  
 KW MHC.  
 FT NON\_TER 1 1  
 FT NON\_TER 81 81  
 SQ SEQUENCE 81 AA; 9405 MW; 073087CE CRC32;

Query Match 100.0%; Score 51; DB 7; Length 81;  
 Best Local Similarity 100.0%; Pred. No. 0.013;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RESLRNLRGY 10  
 DB 66 RESLRNLRGY 75

RESULT 6  
 ID O19533 PRELIMINARY; PRT; 81 AA.  
 AC O19533;  
 DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
 DE MHC CLASS I ANTIGEN HLA-B (FRAGMENT).  
 GN HLA-B.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA CHANDANAYINGYONG D., SIRIKONG M., LUANGTRAKOOL K., SRINAK D.,  
 RA RUNGRONG E., BEJCHANDRA S.;  
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF014779; AAB67817.1; -  
 DR PFAM; PF00129; MHC\_I; 1.

KW MHC.  
FT NON\_TER 1 1  
SQ SEQUENCE 81 AA; 9405 MW; 073087CE CRC32;

Query Match 100.0%; Score 51; DB 7; Length 81;  
Best Local Similarity 100.0%; Pred. No. 0.013;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RESLNLRGY 10  
DB 66 RESLNLRGY 75

RESULT 7  
O19535 PRELIMINARY; PRT; 81 AA.  
AC O19535;  
DT 01-JAN-1998 (TREMELREL. 05, Created)  
DT 01-JAN-1998 (TREMELREL. 05, Last sequence update)  
DT 01-NOV-1998 (TREMELREL. 08, Last annotation update)  
DE MHC CLASS I ANTIGEN HLA-B (FRAGMENT).  
GN HLA-B.

OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA CHANDANAYINGYONG D., SIRIKONG M., LUANGTRAKOOL K., SRINAK D.,  
RA RUNGROUNG E., BEJCHANDRA S.;  
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF014781; AAB67819.1; -  
DR PFAM; PF00129; MHC\_I; 1.  
KW MHC.

FT NON\_TER 1 1  
SQ SEQUENCE 81 AA; 9405 MW; 073087CE CRC32;

Query Match 100.0%; Score 51; DB 7; Length 81;  
Best Local Similarity 100.0%; Pred. No. 0.013;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RESLNLRGY 10  
DB 66 RESLNLRGY 75

RESULT 8  
O19537 PRELIMINARY; PRT; 81 AA.  
AC O19537;  
DT 01-JAN-1998 (TREMELREL. 05, Created)  
DT 01-JAN-1998 (TREMELREL. 05, Last sequence update)  
DT 01-NOV-1998 (TREMELREL. 08, Last annotation update)  
DE MHC CLASS I ANTIGEN HLA-B (FRAGMENT).  
GN HLA-B.

OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA CHANDANAYINGYONG D., SIRIKONG M., LUANGTRAKOOL K., SRINAK D.,  
RA RUNGROUNG E., BEJCHANDRA S.;  
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF014783; AAB67821.1; -  
DR PFAM; PF00129; MHC\_I; 1.  
KW MHC.

FT NON\_TER 1 1  
SQ SEQUENCE 81 AA; 9405 MW; 073087CE CRC32;

Query Match 100.0%; Score 51; DB 7; Length 81;  
Best Local Similarity 100.0%; Pred. No. 0.013;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RESLNLRGY 10  
DB 66 RESLNLRGY 75

RESULT 9  
O19539 PRELIMINARY; PRT; 83 AA.  
AC O19539;  
DT 01-JAN-1998 (TREMELREL. 05, Created)  
DT 01-JAN-1998 (TREMELREL. 05, Last sequence update)  
DT 01-NOV-1998 (TREMELREL. 08, Last annotation update)  
DE MHC CLASS I ANTIGEN HLA-B (FRAGMENT).  
GN HLA-B.

OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA CHANDANAYINGYONG D., SIRIKONG M., LUANGTRAKOOL K., SRINAK D.,  
RA RUNGROUNG E., BEJCHANDRA S.;  
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF014785; AAB67823.1; -  
DR PFAM; PF00129; MHC\_I; 1.  
KW MHC.

FT NON\_TER 1 1  
SQ SEQUENCE 83 AA; 9731 MW; 24B8D666 CRC32;

Query Match 100.0%; Score 51; DB 7; Length 83;  
Best Local Similarity 100.0%; Pred. No. 0.013;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RESLNLRGY 10  
DB 68 RESLNLRGY 77

RESULT 10  
O19541 PRELIMINARY; PRT; 83 AA.  
AC O19541;  
DT 01-JAN-1998 (TREMELREL. 05, Created)  
DT 01-JAN-1998 (TREMELREL. 05, Last sequence update)  
DT 01-NOV-1998 (TREMELREL. 08, Last annotation update)  
DE MHC CLASS I ANTIGEN HLA-B (FRAGMENT).  
GN HLA-B.

OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA CHANDANAYINGYONG D., SIRIKONG M., LUANGTRAKOOL K., SRINAK D.,  
RA RUNGROUNG E., BEJCHANDRA S.;  
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF014787; AAB67825.1; -  
DR PFAM; PF00129; MHC\_I; 1.  
KW MHC.

FT NON\_TER 1 1  
SQ SEQUENCE 83 AA; 9731 MW; 24B8D666 CRC32;

Query Match 100.0%; Score 51; DB 7; Length 83;  
Best Local Similarity 100.0%; Pred. No. 0.013;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RESLRLRGY 10  
| | | | | | | | | |  
DB 68 RESLRLRGY 77

## RESULT 11

O19543  
ID O19543 PRELIMINARY; PRT; 83 AA.  
AC O19543;  
DT 01-JAN-1998 (TREMBlrel. 05, Created)  
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)  
DE MHC CLASS I ANTIGEN HLA-B (FRAGMENT).  
GN HLA-B.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA CHANDANAYONG D., SIRIKONG M., LUANGTRAKOOL K., SRINAK D.,  
RA RUNGROUNG E., BEJCHANDRA S.;  
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF04789; AAB67827.1; -.  
DR PFAM; PF00129; MHC\_I; 1.  
KW MHC.  
FT NON\_TER 1 1  
FT NON\_TER 83 83  
SQ SEQUENCE 83 AA; 9731 MW; 24B8D666 CRC32;

Query Match 100.0%; Score 51; DB 7; Length 83;  
Best Local Similarity 100.0%; Pred. No. 0.013;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RESLRLRGY 10  
| | | | | | | | | |  
DB 68 RESLRLRGY 77

## RESULT 12

Q9Y452  
ID Q9Y452 PRELIMINARY; PRT; 89 AA.  
AC Q9Y452;  
DT 01-NOV-1999 (TREMBlrel. 12, Created)  
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)  
DE HUMAN LEUCOCYTE ANTIGEN B (FRAGMENT).  
GN HLA-B.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA CARTER V., DUNN P.P.;  
RT "Identification of a novel HLA-B\*07 allele."  
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ237594; CAB40714.1; -.  
FT NON\_TER 1 1  
FT NON\_TER 89 89  
SQ SEQUENCE 89 AA; 10609 MW; 659B735B CRC32;

Query Match 100.0%; Score 51; DB 4; Length 89;  
Best Local Similarity 100.0%; Pred. No. 0.014;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RESLRLRGY 10  
| | | | | | | | | |  
DB 74 RESLRLRGY 83

## RESULT 13

P79615

ID P79615 PRELIMINARY; PRT; 89 AA.  
AC P79615;  
DT 01-MAY-1997 (TREMBlrel. 03, Created)  
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)  
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)  
DE MHC CLASS I HLA-B (FRAGMENT).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA CHOPEK M., CAO K., ZHANG G.H.;  
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U88249; AAB48493.1; -.  
DR PFAM; PF00129; MHC\_I; 1.  
KW MHC.  
FT NON\_TER 1 1  
FT NON\_TER 89 89  
SQ SEQUENCE 89 AA; 10489 MW; A7D3DF93 CRC32;

Query Match 100.0%; Score 51; DB 7; Length 89;  
Best Local Similarity 100.0%; Pred. No. 0.014;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RESLRLRGY 10  
| | | | | | | | | |  
DB 74 RESLRLRGY 83

## RESULT 14

P79620  
ID P79620 PRELIMINARY; PRT; 89 AA.  
AC P79620;  
DT 01-MAY-1997 (TREMBlrel. 03, Created)  
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)  
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)  
DE MHC CLASS I HLA-B (FRAGMENT).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA CHOPEK M., CAO K., ZHANG G.H.;  
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U88254; AAB48498.1; -.  
DR PFAM; PF00129; MHC\_I; 1.  
KW MHC.  
FT NON\_TER 1 1  
FT NON\_TER 89 89  
SQ SEQUENCE 89 AA; 10501 MW; FA3EFDCD CRC32;

Query Match 100.0%; Score 51; DB 7; Length 89;  
Best Local Similarity 100.0%; Pred. No. 0.014;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RESLRLRGY 10  
| | | | | | | | | |  
DB 74 RESLRLRGY 83

## RESULT 15

O19548  
ID O19548 PRELIMINARY; PRT; 89 AA.  
AC O19548;  
DT 01-JAN-1998 (TREMBlrel. 05, Created)  
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)  
DE MHC CLASS I ANTIGEN (FRAGMENT).  
GN HLA-B.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;



OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA CAO K., BURDETT L., ZHANG H., FERNANDEZ-VINA M.;  
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF016299; AAB69444.1; -  
DR PFAM; PF00129; MHC\_I; 1.  
KW MHC.  
FT NON\_TER 1  
FT NON\_TER 89  
SQ SEQUENCE 89 AA; 10550 MW; 11452E40 CRC32;

Query Match 100.0%; Score 51; DB 7; Length 89;  
Best Local Similarity 100.0%; Pred. NO. 0.014;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RESLRNLRGY 10  
Db 74 RESLRNLRGY 83

Search completed: February 8, 2000, 19:16:13  
Job time: 21500 sec

**THIS PAGE BLANK (USPTO)**

OM of: US-08-653-294-24 to: GenEmbl.\* out\_format : pfs  
Date: Feb 8, 2000 10:22 PM  
About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlp  
-O=/cgn1.1/USPTO\_spool/US08653294/runat\_04022000\_160701\_15779/app\_query.fasta.2  
-DB=GenEmbl -QPMI=fastcap -SUFFIX=rge -GAPOP=12.000 -GAPEXT=4.500  
-MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -QGAPOP=4.500  
-QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000  
-FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000  
-DELEX=7.000 -START=1 -MATRIX=blossum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR\_SCORE=pct -ALIGN=15 -MODE=LOCAL  
-OUTFWT=pfs -NORM=ext -MINLEN=0 -MAXLEN=100000 -USER=US08653294  
-NCPU=6 -ICPU=3 -NO\_XLPXY -WAIT -THREADS=1

## Search information block:

Query: US-08-653-294-24

Query length: 10

Database: GenEmbl.\*

Database sequences: 821193

Database length: -1518192014

Search time (sec): 10176.920000

## score\_list:

Sequence	Strd Orig	zscore	EScore	Len	Documentation
gb_pr2:HSGLABK1	+	51.00	179.68	243	U91332 Human cell line THAI DCH
gb_pr2:HSGLABM1	+	51.00	179.68	243	U91334 Human cell line THAI DCH
gb_pr2:HSGLABO1	+	51.00	179.68	243	U91335 Human cell line THAI DCH
gb_pr2:AF014769	+	51.00	179.58	246	AF014769 Homo sapiens MHC class
gb_pr2:AF014771	+	51.00	179.58	246	AF014771 Homo sapiens MHC class
gb_pr2:AF014773	+	51.00	179.58	246	AF014773 Homo sapiens MHC class
gb_pr2:AF014775	+	51.00	179.58	246	AF014775 Homo sapiens MHC class
gb_pr2:AF014777	+	51.00	179.58	246	AF014777 Homo sapiens MHC class
gb_pr2:AF014779	+	51.00	179.58	246	AF014779 Homo sapiens MHC class
gb_pr2:AF014781	+	51.00	179.58	246	AF014781 Homo sapiens MHC class
gb_pr2:AF014783	+	51.00	179.58	246	AF014783 Homo sapiens MHC class
gb_pr2:AF014785	+	51.00	179.45	250	AF014785 Homo sapiens MHC class
gb_pr2:AF014787	+	51.00	179.45	250	AF014787 Homo sapiens MHC class
gb_pr2:AF014789	+	51.00	179.45	250	AF014789 Homo sapiens MHC class
gb_pr2:HSB1501N1	+	51.00	179.33	254	US9865 Human MHC class I protei
gb_pr1:D89333	+	51.00	178.85	270	D89333 Homo sapiens DNA for MHC
gb_pr1:HSB131852	+	51.00	178.85	270	AJ131852 Homo sapiens HLA-B gen
gb_pr1:HSB17603	+	51.00	178.85	270	AJ007603 Homo sapiens HLA-B*151
gb_pr1:HSB17605	+	51.00	178.85	270	AJ007605 Homo sapiens HLA-B*151
gb_pr1:UMWHB350S2	+	51.00	178.85	270	M28110 Human MHC class I cell s
gb_pr2:AF002275	+	51.00	178.85	270	AF002275 Homo sapiens MHC class
gb_pr2:AF016299	+	51.00	178.85	270	AF016299 Homo sapiens MHC class
gb_pr2:HS01300S1	+	51.00	178.85	270	U93913 Homo sapiens MHC class I
gb_pr2:HS05602V1	+	51.00	178.85	270	U93913 Homo sapiens MHC class I
gb_pr2:HSB132713	+	51.00	178.85	270	AJ132713 Homo sapiens HLA-B gen
gb_pr2:HSB237594	+	51.00	178.85	270	AJ237594 Homo sapiens HLA-B gen
gb_pr2:HSB243433	+	51.00	178.85	270	AJ243433 Homo sapiens HLA-B gen
gb_pr2:HSB243737	+	51.00	178.85	270	AJ243737 Homo sapiens HLA-B gen
gb_pr2:HSB249317	+	51.00	178.85	270	AJ249317 Homo sapiens HLA-B gen
gb_pr2:HSB0801S1	+	51.00	178.85	270	U67330 Human MHC class I antige
gb_pr2:HSB1505V1	+	51.00	178.85	270	AF017080 Homo sapiens MHC class
gb_pr2:HSB3501V1	+	51.00	178.85	270	U76392 Human HLA-B gene, allele
gb_pr2:HSB3511H1	+	51.00	178.85	270	U76394 Human HLA-B gene, allele
gb_pr2:HSB3901V1	+	51.00	178.85	270	U76394 Human HLA-B gene, allele
gb_pr2:HSB39061S1	+	51.00	178.85	270	U76396 Human HLA-B gene, allele
gb_pr2:HSB40N1	+	51.00	178.85	270	AF002268 Homo sapiens MHC class
gb_pr2:HSB5502V1	+	51.00	178.85	270	AF008926 Homo sapiens MHC class
gb_pr2:HSB560S1	+	51.00	178.85	270	U57746 Human MHC class I antige
gb_pr2:HSB1AB60N1	+	51.00	178.85	270	U80670 Human class I MHC antige
gb_pr2:HSB1ABM1	+	51.00	178.85	270	U90244 Human MHC class I antige
gb_pr2:HSB1ABV01	+	51.00	178.85	270	U55022 Human MHC class I antige
gb_pr2:HSB1ADT1	+	51.00	178.85	270	U91330 Homo sapiens MHC class I

gb\_pr2:HSHYUNAL + 51.00 178.85 0.0613 270 ! U52177 Human MHC class I ant  
gb\_pr2:HSU88249 + 51.00 178.85 0.0613 270 ! U88249 Human MHC class I HLA  
gb\_pr4:HUM31HLA1 + 51.00 178.85 0.0613 270 ! AF181842 Homo sapiens MHC cl

seq\_name: gb\_pr2:HSGLABK1

## seq\_documentation\_block:

LOCUS HSHLABK1 243 bp DNA 26-MAR-1997  
DEFINITION Human cell line THAI DCH022 MHC class I HLA-B gene (allele  
HLA-B\*1521), exon 2.

ACCESSION U91332.1 GI:1906671

VERSION 1 of 2

KEYWORDS

SEGMENT

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 243)

AUTHORS

Chandanayyong,D., Sirikong,M., Longta,K.,

Rungroung,E. and Bejchandra,S.

B15 alleles (B\*1521)

Unpublished

2 (bases 1 to 243)

AUTHORS

Chandanayyong,D., Sirikong,M., Longta,K.,

Rungroung,E. and Bejchandra,S.

Direct Submission

TITLE

Submitted (26-FEB-1997) Transfusion Medicine, Faculty of Medicine,

Siriraj Hospital, Mahidol University, Prannok Road, Bangkok 10700,

Thailand

FEATURES

Location/Qualifiers

1..243

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/map="6p21"

/cell\_line="THAI DCH022"

/cell\_type="lymphoblastoid"

&lt;1..243

/gene="HLA-B"

/number=2

BASE COUNT 53 a 79 c 83 g 28 t

ORIGIN

alignment\_scores:

Quality: 51.00 Length: 10

Ratio: 5.100 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-08-653-294-24 x HSHLABK1 ..

Align seg 1/1 to: HSHLABK1 from: 1 to: 243

1 ArgGluSerLeuArgAsnLeuArgGlyTyr 10

|||||

195 CGAGAGAGCTGCGGACCTGCGGCGCTAC 224

seq\_name: gb\_pr2:HSGLABM1

## seq\_documentation\_block:

LOCUS HSHLABM1 243 bp DNA 26-MAR-1997  
DEFINITION Human cell line THAI DCH023 MHC class I HLA-B gene (allele  
HLA-B\*1525), exon 2.

ACCESSION U91334

VERSION 1 of 2

KEYWORDS

SEGMENT

SOURCE

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Hominidae; Homo.



/protein\_id="AAB67807.1"  
/db\_xref="GI:2345104"  
/translation="TAMSRPGRGEPRTIAGVYDDTQFVRFSDAASPRMAPRAPWIE"  
QEGPEYWRNTQISKNTQTYRESLNLRLGYNQSEA"  
BASE COUNT 55 a 81 c 82 g 28 t  
ORIGIN

alignment\_scores:  
Quality: 51.00 Length: 10  
Ratio: 5.100 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-08-653-294-24 x AF014769 ..  
Align seg 1/1 to: AF014769 from: 1 to: 246

1 ArgGluSerLeuArgAsnLeuArgGlyTyr 10  
|||||  
198 CGAGAGAGCCTGCGAACCTGCGGCGCTAC 227

seq\_name: gb\_pr2:AF014771

seq\_documentation\_block:  
LOCUS AF014771 246 bp DNA PRI 27-AUG-1997  
DEFINITION Homo sapiens MHC class I antigen HLA-B gene (HLA-B\*1502 allele),  
exon 2 and partial cds.

ACCESSION AF014771

VERSION AF014771.1 GI:2345107

KEYWORDS

SOURCE

ORGANISM

human.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 246)

AUTHORS

Chandanayingyong,D., Sirikong,M., Luangtrakool,K., Srinak,D.,

Rungroung,E. and Bejchandra,S.

B15 alleles (B\*1502)

REFERENCE 2 (bases 1 to 246)

AUTHORS

Chandanayingyong,D., Sirikong,M., Luangtrakool,K., Srinak,D.,

Rungroung,E. and Bejchandra,S.

Direct Submission

Submitted (18-JUL-1997) Transfusion Medicine, Faculty of Medicine,

Sirira Hospital, Mahidol University, Prannok Road, Bangkok 10700,

Thailand

FEATURES

source

1..246

Location/Qualifiers

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/cell\_line="Thai DCH4061"

/chromosome="6"

/map="6p21"

/cell\_type="lymphoblastoid"

<1..>246

/gene="HLA-B"

/allele="HLA-B\*1502"

/note="2"

1..246

/gene="HLA-B"

/codon\_start=3

/product="MHC class I antigen HLA-B"

/protein\_id="AAB67807.1"

/db\_xref="GI:2345108"

/translation="TAMSRPGRGEPRTIAGVYDDTQFVRFSDAASPRMAPRAPWIE"

QEGPEYWRNTQISKNTQTYRESLNLRLGYNQSEA"  
BASE COUNT 55 a 81 c 82 g 28 t  
ORIGIN

alignment\_scores:  
Quality: 51.00 Length: 10  
Ratio: 5.100 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-08-653-294-24 x AF014771 ..  
Align seg 1/1 to: AF014771 from: 1 to: 246

1 ArgGluSerLeuArgAsnLeuArgGlyTyr 10  
|||||  
198 CGAGAGAGCCTGCGAACCTGCGGCGCTAC 227

seq\_name: gb\_pr2:AF014773

seq\_documentation\_block:  
LOCUS AF014773 246 bp DNA PRI 27-AUG-1997

DEFINITION Homo sapiens MHC class I antigen HLA-B gene (HLA-B\*1502 allele),  
exon 2 and partial cds.

ACCESSION AF014773

VERSION AF014773.1 GI:2345111

KEYWORDS

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 246)

AUTHORS

Chandanayingyong,D., Sirikong,M., Luangtrakool,K., Srinak,D.,

Rungroung,E. and Bejchandra,S.

B15 alleles (B\*1502)

REFERENCE 2 (bases 1 to 246)

AUTHORS

Chandanayingyong,D., Sirikong,M., Luangtrakool,K., Srinak,D.,

Rungroung,E. and Bejchandra,S.

Direct Submission

Submitted (18-JUL-1997) Transfusion Medicine, Faculty of Medicine,

Sirira Hospital, Mahidol University, Prannok Road, Bangkok 10700,

Thailand

FEATURES

source

1..246

Location/Qualifiers

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/cell\_line="Thai DCH3086"

/chromosome="6"

/map="6p21"

/cell\_type="lymphoblastoid"

<1..>246

/gene="HLA-B"

/allele="HLA-B\*1502"

/note="2"

1..246

/gene="HLA-B"

/codon\_start=3

/product="MHC class I antigen HLA-B"

/protein\_id="AAB67811.1"

/db\_xref="GI:2345112"

/translation="TAMSRPGRGEPRTIAGVYDDTQFVRFSDAASPRMAPRAPWIE"

QEGPEYWRNTQISKNTQTYRESLNLRLGYNQSEA"  
BASE COUNT 55 a 81 c 82 g 28 t  
ORIGIN

alignment\_scores:  
Quality: 51.00 Length: 10  
Ratio: 5.100 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

```

US-08-653-294-24 x AF014773
Align seg 1/1 to: AF014773 from: 1 to: 246

1 ArgGluSerLeuArgAsnLeuArgGlyTyr 10
|||||
198 CGAGAGACCTGCGGACCTGCGGGCTAC 227

seq_name: gb_pr2:AF014775

seq_documentation_block:
LOCUS AF014775 246 bp DNA 27-AUG-1997
DEFINITION Homo sapiens MHC class I antigen HLA-B gene (HLA-B*1502 allele),
exon 2 and partial cds.
ACCESSION AF014775
VERSION AF014775.1 GI:2345115
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 246)
AUTHORS Chandanayingyong,D., Sirikong,M., Luangtrakool,K., Srinak,D.,
Rungroung,E. and Bejchandra,S.
TITLE B15 alleles (B*1502)
JOURNAL Unpublished
REFERENCE
2 (bases 1 to 246)
AUTHORS Chandanayingyong,D., Sirikong,M., Luangtrakool,K., Srinak,D.,
Rungroung,E. and Bejchandra,S.
TITLE Direct Submission
JOURNAL Submitted (18-JUL-1997) Transfusion Medicine, Faculty of Medicine,
Siriraj Hospital, Mahidol University, Prannok Road, Bangkok 10700,
Thailand

FEATURES
source 1..246
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="Thai DCH018"
/chromosome="6"
/map="6p21"
/cell_type="lymphoblastoid"
<1..>246
/gene="HLA-B"
/allele="HLA-B*1502"
/number=2
/notes="2"
<1..>246
/gene="HLA-B"
/codon_start=3
/product="MHC class I antigen HLA-B"
/protein_id="AAB67813.1"
/db_xref="GI:2345116"
/translation="TAMSRPGRGEPRFIAVGVDVDFVRFSDAASPRMAPRPWIE
QEGPEYDNRNTQISKNTQTYRESLNLRGYNQSEA"
BASE COUNT 55 a 81 c 82 g 28 t
ORIGIN

alignment_scores:
Quality: 51.00 Length: 10
Ratio: 5.100 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-24 x AF014775
..
Align seg 1/1 to: AF014775 from: 1 to: 246

1 ArgGluSerLeuArgAsnLeuArgGlyTyr 10
|||||
198 CGAGAGACCTGCGGACCTGCGGGCTAC 227

seq_name: gb_pr2:AF014775

seq_documentation_block:
LOCUS AF014775 246 bp DNA 27-AUG-1997
DEFINITION Homo sapiens MHC class I antigen HLA-B gene (HLA-B*1502 allele),
exon 2 and partial cds.
ACCESSION AF014775
VERSION AF014775.1 GI:2345115
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 246)
AUTHORS Chandanayingyong,D., Sirikong,M., Luangtrakool,K., Srinak,D.,
Rungroung,E. and Bejchandra,S.
TITLE B15 alleles (B*1502)
JOURNAL Unpublished
REFERENCE
2 (bases 1 to 246)
AUTHORS Chandanayingyong,D., Sirikong,M., Luangtrakool,K., Srinak,D.,
Rungroung,E. and Bejchandra,S.
TITLE Direct Submission
JOURNAL Submitted (18-JUL-1997) Transfusion Medicine, Faculty of Medicine,
Siriraj Hospital, Mahidol University, Prannok Road, Bangkok 10700,
Thailand

FEATURES
source 1..246
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="Thai DCH018"
/chromosome="6"
/map="6p21"
/cell_type="lymphoblastoid"
<1..>246
/gene="HLA-B"
/allele="HLA-B*1502"
/number=2
/notes="2"
<1..>246
/gene="HLA-B"
/codon_start=3
/product="MHC class I antigen HLA-B"
/protein_id="AAB67813.1"
/db_xref="GI:2345116"
/translation="TAMSRPGRGEPRFIAVGVDVDFVRFSDAASPRMAPRPWIE
QEGPEYDNRNTQISKNTQTYRESLNLRGYNQSEA"
BASE COUNT 55 a 81 c 82 g 28 t
ORIGIN

alignment_scores:
Quality: 51.00 Length: 10
Ratio: 5.100 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-24 x AF014775
..
Align seg 1/1 to: AF014775 from: 1 to: 246

1 ArgGluSerLeuArgAsnLeuArgGlyTyr 10
|||||
198 CGAGAGACCTGCGGACCTGCGGGCTAC 227

```

```

seq_name: gb_pr2:AF014777

seq_documentation_block:
LOCUS AF014777 246 bp DNA 27-AUG-1997
DEFINITION Homo sapiens MHC class I antigen HLA-B gene (HLA-B*1502 allele),
exon 2 and partial cds.
ACCESSION AF014777
VERSION AF014777.1 GI:2345119
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 246)
AUTHORS Chandanayingyong,D., Sirikong,M., Luangtrakool,K., Srinak,D.,
Rungroung,E. and Bejchandra,S.
TITLE B15 alleles (B*1502)
JOURNAL Unpublished
REFERENCE
2 (bases 1 to 246)
AUTHORS Chandanayingyong,D., Sirikong,M., Luangtrakool,K., Srinak,D.,
Rungroung,E. and Bejchandra,S.
TITLE Direct Submission
JOURNAL Submitted (18-JUL-1997) Transfusion Medicine, Faculty of Medicine,
Siriraj Hospital, Mahidol University, Prannok Road, Bangkok 10700,
Thailand

FEATURES
Location/Qualifiers
source 1..246
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="Thai DCH017"
/chromosome="6"
/map="6p21"
/cell_type="lymphoblastoid"
<1..>246
/gene="HLA-B"
/allele="HLA-B*1502"
/number=2
/notes="2"
<1..>246
/gene="HLA-B"
/codon_start=3
/product="MHC class I antigen HLA-B"
/protein_id="AAB67815.1"
/db_xref="GI:2345120"
/translation="TAMSRPGRGEPRFIAVGVDVDFVRFSDAASPRMAPRPWIE
QEGPEYDNRNTQISKNTQTYRESLNLRGYNQSEA"
BASE COUNT 55 a 81 c 82 g 28 t
ORIGIN

alignment_scores:
Quality: 51.00 Length: 10
Ratio: 5.100 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-24 x AF014777
..
Align seg 1/1 to: AF014777 from: 1 to: 246

1 ArgGluSerLeuArgAsnLeuArgGlyTyr 10
|||||
198 CGAGAGACCTGCGGACCTGCGGGCTAC 227

seq_name: gb_pr2:AF014779

seq_documentation_block:
LOCUS AF014779 246 bp DNA 27-AUG-1997
DEFINITION Homo sapiens MHC class I antigen HLA-B gene (HLA-B*1502 allele),
exon 2 and partial cds.

```

```

ACCESSION      AF014779
VERSION         AF014779.1  GI:2345123
KEYWORDS
SOURCE          human.
ORGANISM        Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
                Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS        Chandanayingyong,D., Sirikong,M., Luangtrakool,K., Srinak,D.,
                Rungroung,E. and Bejchandra,S.
TITLE          1 (bases 1 to 246)
JOURNAL        Chananayingyong,D., Sirikong,M., Luangtrakool,K., Srinak,D.,
                Rungroung,E. and Bejchandra,S.
AUTHORS        B15 alleles (B*1502)
                Unpublished
REFERENCE      2 (bases 1 to 246)
                Chandanayingyong,D., Sirikong,M., Luangtrakool,K., Srinak,D.,
                Rungroung,E. and Bejchandra,S.
TITLE          Direct Submission
JOURNAL        Submitted (18-JUL-1997) Transfusion Medicine, Faculty of Medicine,
                Siriraj Hospital, Mahidol University, Prannok Road, Bangkok 10700,
                Thailand
FEATURES
source         Location/Qualifiers
                1..246
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /cell_line="Thai DCH002"
                /chromosome="6"
                /map="p21"
                /cell_type="lymphoblastoid"
                <1..>246
                /gene="HLA-B"
                /allele="HLA-B*1502"
                1..246
                /gene="HLA-B"
                /note="2"
                /number=2
                <1..>246
                /gene="HLA-B"
                /codon_start=3
                /product="MHC class I antigen HLA-B"
                /protein_id="AAB67817.1"
                /db_xref="GI:2345124"
                /translation="TMSRPGRGEPRTFANGVDDTQFVRDSDAASPRMAPRPWIE
                QEGPEYDWRNTQISKNTQTYRESLRNLRGYNQSEA"
BASE COUNT     55 a 81 c 82 g 28 t
ORIGIN
alignment_scores:
    Quality: 51.00      Length: 10
    Ratio: 5.100       Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 100.000
alignment_block:
US-08-653-294-24 x AF014779 ..
Align seg 1/1 to: AF014779 from: 1 to: 246
1 ArgGluserLeuArgAsnLeuArgGlyTyr 10
|||||
198 CGAGAGAGCGCTGCGGAACCTGCGGCGCTAC 227
seq_name: gb_pr2:AF014781
seq_documentation_block:
LOCUS      AF014781      246 bp      DNA      PRI      27-AUG-1997
DEFINITION Homo sapiens MHC class I antigen HLA-B gene (HLA-B*1502 allele),
            exon 2 and partial cds.
ACCESSION  AF014781
VERSION    AF014781.1  GI:2345131
KEYWORDS
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 246)
            Chandanayingyong,D., Sirikong,M., Luangtrakool,K., Srinak,D.,
            Rungroung,E. and Bejchandra,S.
AUTHORS    B15 alleles (B*1502)
            Unpublished
REFERENCE  2 (bases 1 to 246)
            Chandanayingyong,D., Sirikong,M., Luangtrakool,K., Srinak,D.,
            Rungroung,E. and Bejchandra,S.
TITLE
JOURNAL
AUTHORS

```

Rungroung,E. and Bejchandra,S.  
 Direct Submission  
 Submitted (18-JUL-1997) Transfusion Medicine, Faculty of Medicine,  
 Siriraj Hospital, Mahidol University, Prannok Road, Bangkok 10700,  
 Thailand

# FEATURES

source  
 Location/Qualifiers  
 1. .246

/organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /cell\_line="Thai DCH016"  
 /chromosome="6"  
 /map="6p21"  
 /cell\_type="lymphoblastoid"  
 <1. .>246

## gene

/gene="HLA-B"  
 /allele="HLA-B\*1502"  
 1. .246

## exon

/gene="HLA-B"  
 /note="2"  
 /number=2  
 <1. .>246

## CDS

/gene="HLA-B"  
 /codon\_start=3  
 /product="MHC class I antigen HLA-B"  
 /protein\_id="AAB67821.1"  
 /db\_xref="GI:2345132"  
 /translation="TAMSRPGCEPRFIAGVYDDTQFVRFDSDAASPRMAPRWIE  
 QEGPEYDRETOISKTNTQTYRESLNLRGYNQSEA"

BASE COUNT 55 a 81 c 82 g 28 t  
 ORIGIN

## alignment\_scores:

Quality: 51.00 Length: 10  
 Ratio: 5.100 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-08-653-294-24 x AF014783 ..

Align seg 1/1 to: AF014783 from: 1 to: 246

1 ArgGlusSerLeuArgAsnLeuArgGlyTyr 10  
 |||||

198 CGAGAGAGCTCGGACCTGCGGGCTAC 227

seq\_name: gb\_pr2:AF014785

## seq\_documentation\_block:

LOCUS AF014785 250 bp DNA PRI 27-AUG-1997  
 DEFINITION Homo sapiens MHC class I antigen HLA-B gene (HLA-B\*1525 allele),  
 exon 2 and partial cds.

ACCESSION AF014785

VERSION AF014785.1 GI:2345135

## KEYWORDS

SOURCE human.

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 250)

AUTHORS Chananayong,D., Sirikong,M., Luangtrakool,K., Srinak,D.,  
 Rungroung,E. and Bejchandra,S.

TITLE B15 alleles (B\*1525)

REFERENCE 2 (bases 1 to 250)

AUTHORS Chananayong,D., Sirikong,M., Luangtrakool,K., Srinak,D.,  
 Rungroung,E. and Bejchandra,S.

TITLE Direct Submission

JOURNAL Submitted (18-JUL-1997) Transfusion Medicine, Faculty of Medicine,  
 Siriraj Hospital, Mahidol University, Prannok Road, Bangkok 10700,  
 Thailand

## FEATURES

source

Location/Qualifiers  
 1. .250

/organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /cell\_line="Thai DCH3258"  
 /chromosome="6"  
 /map="6p21"  
 /cell\_type="lymphoblastoid"

## gene

1. .250

/organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /cell\_line="Thai DCH1109"  
 /chromosome="6"  
 /map="6p21"  
 /cell\_type="lymphoblastoid"  
 1. .250

## gene

/gene="HLA-B"  
 /allele="HLA-B\*1525"  
 1. .250

## exon

/gene="HLA-B"  
 /note="2"  
 /number=2  
 <1. .>250

## CDS

/gene="HLA-B"  
 /codon\_start=1  
 /product="MHC class I antigen HLA-B"  
 /protein\_id="AAB67823.1"  
 /db\_xref="GI:2345136"  
 /translation="FYTAMSRPGCEPRFIAGVYDDTQFVRFDSDAASPRMAPRAPW  
 IEQSGPEYWDRETOISKTNTQTYRESLNLRGYNQSEA"

BASE COUNT 54 a 81 c 84 g 31 t  
 ORIGIN

## alignment\_scores:

Quality: 51.00 Length: 10  
 Ratio: 5.100 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-08-653-294-24 x AF014785 ..

Align seg 1/1 to: AF014785 from: 1 to: 250

1 ArgGlusSerLeuArgAsnLeuArgGlyTyr 10  
 |||||

202 CGAGAGAGCTCGGACCTGCGGGCTAC 231

seq\_name: gb\_pr2:AF014787

## seq\_documentation\_block:

LOCUS AF014787 250 bp DNA PRI 27-AUG-1997  
 DEFINITION Homo sapiens MHC class I antigen HLA-B gene (HLA-B\*1525 allele),  
 exon 2 and partial cds.

ACCESSION AF014787

VERSION AF014787.1 GI:2345139

## KEYWORDS

SOURCE human.

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 250)

AUTHORS Chananayong,D., Sirikong,M., Luangtrakool,K., Srinak,D.,  
 Rungroung,E. and Bejchandra,S.

TITLE B15 alleles (B\*1525)

REFERENCE 2 (bases 1 to 250)

AUTHORS Chananayong,D., Sirikong,M., Luangtrakool,K., Srinak,D.,  
 Rungroung,E. and Bejchandra,S.

TITLE Direct Submission

JOURNAL Submitted (18-JUL-1997) Transfusion Medicine, Faculty of Medicine,  
 Siriraj Hospital, Mahidol University, Prannok Road, Bangkok 10700,  
 Thailand

## FEATURES

source

Location/Qualifiers  
 1. .250

/organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /cell\_line="Thai DCH3258"  
 /chromosome="6"  
 /map="6p21"  
 /cell\_type="lymphoblastoid"

## gene

1. .250



```
/gene="HLA-B"
/allele="HLA-B*1525"
1..250
/exon
```

```
/gene="HLA-B"
/note="2"
/number=2
<1..>250
CDS
```

```
/gene="HLA-B"
/codon_start=1
/product="MHC class I antigen HLA-B"
/protein_id="AAB67825.1"
/db_xref="GI:2345140"
/translation="FYTAMSRPGRGEPFRFIAGVYDDTQFVRFSDAASPRMAPRAPW
IEQEGPYWDRETOISKNTQTYRESLRLRGYNQSEA"
```

```
BASE COUNT      54 a      81 c      84 g      31 t
ORIGIN
```

```
alignment_scores:
  Quality: 51.00      Length: 10
  Ratio: 5.100      Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
```

```
alignment_block:
US-08-653-294-24 x AF014787 ..
```

```
Align seg 1/1 to: AF014787 from: 1 to: 250
```

```
1 ArgGluSerLeuArgAsnLeuArgGlyTyr 10
|||||
202 CGAGAGAGCCTGCGGAACCTGCGGCGCTAC 231
```

```
seq_name: gb_pr2:AF014789
```

```
seq_documentation_block:
LOCUS AF014789 250 bp DNA PRI 27-AUG-1997
DEFINITION Homo sapiens MHC class I antigen HLA-B gene (HLA-B*1525 allele),
exon 2 and partial cds.
```

```
ACCESSION AF014789
VERSION AF014789.1 GI:2345143
```

```
KEYWORDS
SOURCE human.
```

```
ORGANISM Homo sapiens
```

```
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

```
REFERENCE 1 (bases 1 to 250)
```

```
AUTHORS Chandanayingyong,D., Sirikong,M., Luangtrakool,K., Srinak,D.,
Rungrong,E. and Bejchandra,S.
```

```
TITLE B15 alleles (B*1525)
```

```
JOURNAL Unpublished
```

```
REFERENCE 2 (bases 1 to 250)
```

```
AUTHORS Chandanayingyong,D., Sirikong,M., Luangtrakool,K., Srinak,D.,
Rungrong,E. and Bejchandra,S.
```

```
TITLE Direct Submission
```

```
JOURNAL Submitted (18-JUL-1997) Transfusion Medicine, Faculty of Medicine,
Siriraj Hospital, Mahidol University, Prannok Road, Bangkok 10700,
Thailand
```

```
FEATURES Location/Qualifiers
```

```
source
1..250
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="Thai DCH025"
/chromosome="6"
/map="6p21"
/cell_type="lymphoblastoid"
```

```
1..250
```

```
gene
/gene="HLA-B"
/allele="HLA-B*1525"
```

```
1..250
```

```
exon
/gene="HLA-B"
```

```
/note="2"
```

```
/number=2
```

```
<1..>250
CDS
```

```
/gene="HLA-B"
/codon_start=1
/product="MHC class I antigen HLA-B"
/protein_id="AAB67827.1"
/db_xref="GI:2345144"
/translation="FYTAMSRPGRGEPFRFIAGVYDDTQFVRFSDAASPRMAPRAPW
IEQEGPYWDRETOISKNTQTYRESLRLRGYNQSEA"
```

```
BASE COUNT      54 a      81 c      84 g      31 t
ORIGIN
```

```
alignment_scores:
  Quality: 51.00      Length: 10
  Ratio: 5.100      Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
```

```
alignment_block:
US-08-653-294-24 x AF014789 ..
```

```
Align seg 1/1 to: AF014789 from: 1 to: 250
```

```
1 ArgGluSerLeuArgAsnLeuArgGlyTyr 10
|||||
202 CGAGAGAGCCTGCGGAACCTGCGGCGCTAC 231
```

```
seq_name: gb_pr2:HSB1501N1
```

```
seq_documentation_block:
LOCUS HSB1501N1 254 bp DNA PRI 10-JUL-1996
DEFINITION Human MHC class I protein HLA-B heavy chain (B*1501new allele)
gene, partial exon 2.
```

```
ACCESSION U59965
```

```
VERSION U59965.1 GI:1399896
```

```
KEYWORDS HLA B antigen; major histocompatibility complex.
```

```
SEGMENT 1 of 2
```

```
SOURCE human.
```

```
ORGANISM Homo sapiens
```

```
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

```
REFERENCE 1 (bases 1 to 254)
```

```
AUTHORS Mitsuishi,Y.
```

```
TITLE Direct Submission
```

```
JOURNAL Submitted (04-JUN-1996) Tissue Typing Laboratory, UCLA School of
Medicine, 950 Veteran Ave., Los Angeles, CA 90095, USA
```

```
FEATURES Location/Qualifiers
```

```
source
```

```
1..254
/organism="Homo sapiens"
```

```
/db_xref="taxon:9606"
```

```
/cell_type="peripheral blood mononuclear cells"
```

```
/chromosome="6"
```

```
/map="6p21.3"
```

```
<1..254
```

```
/gene="HLA-B"
```

```
/number=2
```

```
BASE COUNT      55 a      81 c      85 g      33 t
```

```
ORIGIN
```

```
alignment_scores:
  Quality: 51.00      Length: 10
  Ratio: 5.100      Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
```

```
alignment_block:
US-08-653-294-24 x HSB1501N1 ..
```

```
Align seg 1/1 to: HSB1501N1 from: 1 to: 254
```

```
1 ArgGluSerLeuArgAsnLeuArgGlyTyr 10
|||||
```

```
206 CGAGAGAGCCTGCGGAACCTGCGGCGCTAC 235
```

**THIS PAGE BLANK (USPTO)**

OM of: US-08-653-294-24 to: N\_Geneseq\_36:\* out\_format : pfs  
Date: Feb 8, 2000 7:31 PM  
About: Results were produced by the GenCore software, version 4.5.  
Copyright (c) 1993-2000 CompuGen Ltd.

## Command line parameters:

-MODEL=framet.p2n.model -DEV=xlp  
-Q=/cgnl1/USFTO\_spool/US08653294/runat\_04022000\_160701\_15807/app\_query.fasta.2  
-DB=N\_Geneseq\_36 -QFMT=fastap -SURFIX=ring -GAPOP=12.000  
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000  
-GAPOP=4.500 -GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500  
-GAPEXT=6.000 -GAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500  
-DELOP=6.000 -DEEXT=7.000 -START=1 -MATRIX=blotsum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR\_SCORE=pct  
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0  
-MAXLEN=1000000 -USER=US08653294 -NCPU=6 -ICPU=3 -NO\_XLPXY -WAIT  
-THREADS=1

## Search information block:

Query: US-08-653-294-24  
Query length: 10  
Database: N\_Geneseq\_36:\*  
Database sequences: 311585  
Database length: 125096042  
Search time (sec): 873.190000

## score\_list:

Sequence	Strid	Orig	zScore	Escore	Len	Documentation
N_Geneseq_36:Q12115	+	51.00	169.52	0.0767	1089	HLA-B35 exon. HLA-B35 gene -
N_Geneseq_36:Q75974	+	51.00	157.69	0.3497	4059	HLA-B7 expression vector. New
N_Geneseq_36:Q75973	+	51.00	155.88	0.4411	4965	HLA-B7/beta-2 microglobulin
N_Geneseq_36:Q12116	+	44.00	143.58	2.14	1101	HLA-C exon Cb-1. HLA-C gene, D
N_Geneseq_36:Q12117	+	44.00	143.58	2.14	1101	HLA-C exon Cb-2. HLA-C gene, D
N_Geneseq_36:X22446	+	39.00	123.46	28.18	1323	Human CPM2 DNA. New chemoprot
N_Geneseq_36:X22450	+	39.00	123.46	28.21	1324	Human CPM2 mutant DNA. New ch
N_Geneseq_36:T32217	+	38.00	118.70	51.91	1490	Acyl-ACP-thioesterase coding s
N_Geneseq_36:T32447	+	38.00	117.74	58.75	1659	Rat truncated RSK3 coding sequ
N_Geneseq_36:T35448	+	38.00	117.65	59.37	1674	Truncated human hsk3 coding se
N_Geneseq_36:T35472	+	38.00	115.10	82.38	2224	Rat RSK3 coding sequence. DNA
N_Geneseq_36:T35473	+	38.00	114.19	92.63	2462	Human hsk3 coding sequence. DN
N_Geneseq_36:X08414	+	38.00	113.97	95.20	2521	Human small conductance calciu
N_Geneseq_36:X41493	+	37.00	126.48	19.13	416	Human secreted protein 5' EST
N_Geneseq_36:T30458	+	37.00	76.32	1.2e+04	110000	Rhizobium species plasmid p
N_Geneseq_36:T30459	+	37.00	76.32	1.2e+04	110000	Rhizobium species symbiotic
N_Geneseq_36:Q11953	+	35.00	120.86	39.35	342	Anti-human Rhd BRAD-3 MAb (VH
N_Geneseq_36:Q11953	+	35.00	120.76	39.88	346	Intron 2 from human follistatin
N_Geneseq_36:Q56202	+	35.00	105.64	277.17	1858	Partial coding sequence of sal
N_Geneseq_36:Q50126	+	35.00	105.64	277.17	1858	Plasmid pCK1 SABP cDNA insert.
N_Geneseq_36:T38418	+	35.00	105.64	277.17	1858	pCK1 cDNA insert encoding sal
N_Geneseq_36:Q94624	+	35.00	99.84	583.35	3342	Brush-1 cDNA probe. New cDNA s
N_Geneseq_36:T15158	+	35.00	98.12	726.84	4286	CDK41 cyclin-dependent kinase
N_Geneseq_36:Q34625	+	35.00	97.97	741.13	4359	Brush-1 cDNA. New cDNA sequen
N_Geneseq_36:X33449	+	35.00	97.44	793.74	4626	Oryza sativa L. pistil specifi
N_Geneseq_36:T59528	+	34.00	113.84	96.79	495	Human secreted protein gene 18
N_Geneseq_36:N91608	+	34.00	107.13	228.91	1044	Insert of pCCA10 encoding carb
N_Geneseq_36:Q52673	+	34.00	107.57	279.97	1242	HTLE receptor DNA. Stably co-t
N_Geneseq_36:Q47076	+	34.00	104.68	313.43	1371	B. canis 21B4/rhoxyr antigen
N_Geneseq_36:V01731	+	34.00	100.81	514.80	2108	Mouse liver cancer-originated
N_Geneseq_36:Q43711	+	34.00	99.41	616.03	2463	Sequence of the human gene 5-H
N_Geneseq_36:X13139	+	34.00	85.96	3.5e+03	10996	Enterococcus faecalis genome
N_Geneseq_36:X13401	+	34.00	85.93	3.5e+03	10936	Enterococcus faecalis genome
N_Geneseq_36:T21210	+	34.00	70.94	2.3e+04	58407	Methanococcus jannaschii larg
N_Geneseq_36:X20248	+	34.00	65.25	4.7e+04	110000	Continuation (8 of 10) of
N_Geneseq_36:T94815	+	33.00	139.96	3.40	18	Human leukocyte antigen class I
N_Geneseq_36:T42475	+	33.00	123.93	26.56	107	Hunt1 coding sequence. Hunt1 an
N_Geneseq_36:Q72980	+	33.00	120.03	43.77	165	Huntingtin gene antisense produ
N_Geneseq_36:X23425	+	33.00	107.02	232.16	745	Mouse TRNL3 DNA. New Tumor Necr
N_Geneseq_36:T69757	+	33.00	106.47	249.05	701	Isopentenyl transferase ipt gen
N_Geneseq_36:V18599	+	33.00	102.43	418.34	1168	Mus musculus tumour necrosis f

N\_Geneseq\_36:V43032 + 33.00 101.43 475.43 1305 ! Streptococcus pneumoniae po  
N\_Geneseq\_36:V32620 - 33.00 101.11 495.66 1353 ! Genomic DNA sequence of str  
N\_Geneseq\_36:Q80003 - 33.00 98.73 672.62 1763 ! ADP-glucose-pyrophosphoryla  
N\_Geneseq\_36:V29686 + 33.00 95.40 1.0e+03 2553 ! A. thermophilum thermostabi  
seq\_name: N\_Geneseq\_36:Q12115  
seq\_documentation\_block:  
ID Q12115 standard; DNA; 1089 BP.  
AC Q12115;  
DT 29-AUG-1991 (first entry)  
DE HLA-B35 exon.  
KW Human leukocyte antigen; probe; major histocompatibility complex;  
KW MHC; class I; ss.  
OS Homo sapiens.  
FH Key Location/Qualifiers  
FT cds 1..1089  
FT /tag= a  
PN J03112486-A.  
PD 14-MAY-1991.  
PF 22-SEP-1989; 247697.  
PR 22-SEP-1989; JP-247697.  
PA (OLYU ) OLYMPUS OPTICAL KK.  
DR WPI; 91-182991/25.  
DR P-PSDB; R12464.  
PT HLA-B35 gene - used in DNA probe and transformant cells for  
PT immunising animals, for developing monoclonal antibody.  
PS Claim 1; Page 1; lipp; Japanese.  
CC Probes comprising part of the sequence can be used to identify  
CC class I genes. The DNA can be expressed for immunisation of  
CC animals and prodn. of monoclonal antibodies specific for the  
CC HLA-B35 antigen. See also J03112485 and J03112487.  
SQ Sequence 1089 BP; 221 A; 336 C; 359 G; 173 T;

## alignment\_scores:

Quality: 51.00 Length: 51 10  
Ratio: 5.100 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-08-653-294-24 x Q12115 ..  
Align seg 1/1 to: Q12115 from: 1 to: 1089  
1 ArgGluSerLeuArgAsnLeuArgGlyTyr 10  
295 CGAGAGAGCCTCGGAACCTCGCGCTAC 324

## seq\_name: N\_Geneseq\_36:Q75974

## seq\_documentation\_block:

ID Q75974 standard; cDNA; 4059 BP.  
AC Q75974;  
DT 23-AUG-1995 (first entry)  
DE PHLA-B7 expression vector.  
DE expression vector; PHLA-B7; heavy human HLA-B7; bicistronic mRNA;  
KW light beta-2 microglobulin; class I major histocompatibility complex;  
KW MHC; human leukocyte antigen; HLA; covalently closed circular DNA; ds.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT misc\_feature 1..354  
FT /tag= a  
FT /note= "PBR322 backbone contg. bacterial origin of  
FT replication"  
FT 355..1170  
FT /tag= b  
FT /note= "kanamycin resistance gene open reading frame;  
FT the gene is taken from the transposable element  
FT Tn903"  
FT polyA\_signal complement (1410..1177)  
FT /tag= c  
FT /note= "SV40 polyA signal sequence"

FT intron complement (1412..1560)  
 FT /tag= d  
 FT /note= "SV40 small t intron"  
 FT complement (1561..1794)  
 FT /tag= e  
 FT /note= "3' untranslated region of HLA-B7  
 FT heavy chain mRNA"  
 FT complement (1795..2880)  
 FT /tag= f  
 FT /note= "HLA-B7 open reading frame"  
 FT complement (2886..3415)  
 FT /tag= g  
 FT /note= "Rous sarcoma virus 3' LTR promoter region"  
 FT 3416..4059  
 FT /tag= h  
 FT /note= "pBR322 backbone"  
 FT W09429469-A.  
 PD 22-DEC-1994.  
 PF 27-MAY-1994; U06069.  
 PR 07-JUN-1993; US-074344.  
 PA (UNMI ) UNIV MICHIGAN.  
 PA (VICA-) VICAL INC.  
 PI Lew D, Marquet M, Nabel EG, Nabel GJ;  
 DR WPI; 95-036494/05.  
 PT New vectors for gene therapy, partic for tumours - comprising  
 PT genetic material encoding one or more cistron(s) which express  
 PT immunogenic or therapeutic peptide(s)  
 PS Claim 9; Page 42-43; 50pp; English.  
 CC This HLA-B7 antigen encoding plasmid was developed to incorporate many  
 CC advantageous features, eg. the kanamycin resistance gene. The  
 CC eradication of two open reading frames encoding portions of SV40 viral  
 CC proteins lowers the risk of tumorigenicity. The vector may also operate  
 CC as a cassette into which cistrons may be inserted and removed at will  
 CC for the transcription and subsequent translation of peptides of interest.  
 CC The vector is used partic. for the treatment of neoplastic disease,  
 CC eg. melanoma, and provides enhanced gene delivery and expression  
 CC in vivo.  
 SQ Sequence 4059 BP; 975 A; 1051 C; 1033 G; 1000 T;

alignment\_scores:  
 Quality: 51.00 Length: 10  
 Ratio: 5.100 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
 US-08-653-294-24 x Q75974/rev ..  
 Align seg 1/1 to reverse of: Q75974 from: 1 to: 4059

1 ArgGluSerLeuArgAsnLeuArgGlyTyr 10  
 |||||  
 2589 CGAGAGACCTGCGGACCTGCGGGCTAC 2560

seq\_name: N\_Geneseq\_36:Q75973  
 seq\_documentation\_block:  
 ID Q75973 standard; cDNA; 4965 BP.  
 AC Q75973;  
 DT 23-AUG-1995 (first entry)  
 DE pHLA-B7/beta-2 microglobulin expression vector.  
 KW expression vector; pHLA-B7/beta-2 microglobulin; heavy human HLA-B7;  
 KW light beta-2 microglobulin; class I major histocompatibility complex;  
 KW MHC; bicistronic mRNA; human leukocyte antigen; HLA;  
 KW covalently closed circular DNA; ds.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT 1..529  
 FT /tag= a  
 FT /note= "Rous sarcoma virus LTR promoter domain, derived  
 FT for the Schmidt-Rupin strain nucleotides  
 FT 8673-9146. This region also includes a 56 bp  
 FT region of a synthetic oligonucleotide which

FT modifies this regulatory sequence to effect a  
 FT higher level of expression of downstream  
 FT sequences. The oligonucleotide removes a  
 FT polyadenylation signal sequence originally found  
 FT in the RSV DNA sequence."  
 FT 531..534  
 FT /tag= b  
 FT /label= consensus\_kozak\_signal\_sequence  
 FT 535..1620  
 FT /tag= c  
 FT /note= "HLA-B7 heavy chain open reading frame"  
 FT 535..606  
 FT /tag= d  
 FT /note= "encodes putative signal peptide of the  
 FT HLA-B7 heavy chain"  
 FT 607..1620  
 FT /tag= e  
 FT /note= "encodes putative HLA-B7 heavy chain mature  
 FT peptide"  
 FT 1621..1853  
 FT /tag= f  
 FT /note= "3' untranslated sequence of HLA-B7 heavy  
 FT chain mRNA"  
 FT 1854..1888  
 FT /tag= g  
 FT /note= "Multiple cloning site, forms a junction  
 FT between the HLA-B7 sequence and the EMCV-CITE  
 FT sequence, and is used to facilitate subcloning"  
 FT 1889..2479  
 FT /tag= h  
 FT /note= "murine encephalomyocarditis CAP-independent  
 FT translational enhancer (EMCV-CITE); taken from  
 FT nucleotides 255-843 of cloned EMCV genomic DNA.  
 FT It is a non-coding regulatory sequence functioning  
 FT as an internal entry point for the eukaryotic  
 FT ribosomal subunits when located within a mRNA  
 FT mol. . It enables the translational start codon of  
 FT the beta-2 microglobulin, downstream of the HLA-B7  
 FT stop codon on this bicistronic mRNA to be  
 FT recognised by the ribosome"  
 FT 2480..2839  
 FT /tag= i  
 FT /note= "encodes beta-2 microglobulin; this cDNA is  
 FT deriv. from chimpanzee (differs to the human  
 FT cDNA by only 4 bases)"  
 FT 2840..2846  
 FT /tag= j  
 FT /note= "3' untranslated region of the beta-2  
 FT microglobulin mRNA"  
 FT 2847..2870  
 FT /tag= k  
 FT /note= "synthetic linker"  
 FT 3'UTR .3111  
 FT /tag= l  
 FT /note= "Bovine growth hormone 3'UTR and transcriptional  
 FT terminator; it starts at a blunt-ended BglI site  
 FT within the 3'UTR of the mRNA coding sequence"  
 FT 2979..2984  
 FT /tag= m  
 FT 3112..3151  
 FT /tag= n  
 FT /note= "Synthetic linker to facilitate cloning"  
 FT complement (3151..3967)  
 FT /tag= o  
 FT /note= "Kanamycin resistance gene open reading frame;  
 FT the gene is taken from the transposable element  
 FT Tn903"  
 FT 4014..4965  
 FT /tag= p  
 FT /note= "pBR322 backbone contg. bacterial origin of  
 FT replication, it represents nucleotides 2244-3193"  
 FT W09429469-A.  
 PD 22-DEC-1994.

PF 27-MAY-1994; U06069.  
PR 07-JUN-1993; US-074344.  
PA (UNMI ) UNIV MICHIGAN.  
PA (VICA-) VICAL INC.  
PI Lew D, Marquet M, Nabel EG, Nabel GJ;  
DR WPI; 95-036494/05.  
PT New vectors for gene therapy, partic for tumours - comprising  
PT genetic material encoding one or more cistron(s) which express  
PT immunogenic or therapeutic peptide(s)  
PS Claim 8; Page 41-42; 50pp; English.  
CC The pHLA-B7/beta-2 microglobulin plasmid expression vector, in addition  
CC to the kanamycin resistance gene, contains the plasmid DNA encoding the  
CC heavy (human HLA-B7) and light (beta-2 microglobulin) proteins of a  
CC class I major histocompatibility complex (MHC) antigen. The plasmid is  
CC designed to express these two proteins via a bicistronic mRNA in  
CC eukaryotic cells. Initiation of transcription of the mRNA is dependent  
CC on a Rous Sarcoma Virus (RSV) promoter sequence deriv. from the 3' long  
CC terminal repeat. Termination of transcription is dependent upon the  
CC polyadenylation signal sequence deriv. from the bovine growth hormone  
CC gene. Eukaryotic cell translation of the heavy chain is regulated by the  
CC 5' cap-dependent protein start site. Translation of the light chain is  
CC controlled by the CITE. Finally the replication of the plasmid in  
CC bacterial cells is controlled by the presence of a bacterial origin of  
CC replication. The vector is used partic. for the treatment of neoplastic  
CC disease, eg. melanoma, and provides enhanced gene delivery and expression  
CC in vivo.  
SQ Sequence 4965 BP; 1171 A; 1293 C; 1338 G; 1163 T;

alignment\_scores:  
Quality: 51.00 Length: 10  
Ratio: 5.100 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-08-653-294-24 x Q75973

Align seg 1/1 to: Q75973 from: 1 to: 4965

1 ArgGluSerLeuArgAsnLeuArgGlyTyr 10  
|||||  
829 CGAGAGAGCGCTGCGGAACCTGCGCGGTAC 858

seq\_name: N\_Geneseq\_36:Q12116

seq\_documentation\_block:

ID Q12116 standard; DNA; 1101 BP.

AC Q12116;  
DT 29-AUG-1991 (first entry)  
DE HLA-C exon Cb-1.  
KW Human leukocyte antigen; probe; major histocompatibility complex;  
KW MHC; class I; ss.  
OS Homo sapiens.

FT Key Location/Qualifiers

FT cds 1..1101  
/\*tag= a

PN J03112485-A.

PD 14-MAY-1991.

PF 22-SEP-1989; 247695.

PR 22-SEP-1989; JP-247695.

PA (OLYU ) OLYMPUS OPTICAL KK.

DR WPI; 91-182989/25.

DR P-PSDB; R12465.

PT HLA-C gene, DNA probe and transformant cells - for immunisation

PT of animals and monoclonal antibody development.

PS Claim 1; Page 1; 13pp; Japanese.

CC Probes comprising part of the sequence can be used to identify

CC Class I genes. The DNA can be expressed for immunisation of

CC animals and prodn. of monoclonal antibodies specific for the HLA-C

CC antigen. See also Q12117 (same patent) and J03112486 and J03112487.

CC Sequence 1101 BP; 211 A; 337 C; 377 G; 176 T;

SQ

alignment\_scores:  
Quality: 44.00 Length: 10  
Ratio: 4.889 Gaps: 0  
Percent Similarity: 90.000 Percent Identity: 90.000

alignment\_block:

US-08-653-294-24 x Q12116

Align seg 1/1 to: Q12116 from: 1 to: 1101

1 ArgGluSerLeuArgAsnLeuArgGlyTyr 10  
|||||  
295 CGAGTGAGCGCTGCGGAACCTGCGCGGTAC 324

seq\_name: N\_Geneseq\_36:Q12117

seq\_documentation\_block:

ID Q12117 standard; DNA; 1101 BP.

AC Q12117;

DT 29-AUG-1991 (first entry)

DE HLA-C exon Cb-2.

KW Human leukocyte antigen; probe; major histocompatibility complex;

KW MHC; class I; ss.

OS Homo sapiens.

FT Key Location/Qualifiers

FT cds 1..1101

/\*tag= a

PN J03112485-A.

PD 14-MAY-1991.

PF 22-SEP-1989; 247695.

PR 22-SEP-1989; JP-247695.

PA (OLYU ) OLYMPUS OPTICAL KK.

DR WPI; 91-182989/25.

DR P-PSDB; R12466.

PT HLA-C gene, DNA probe and transformant cells - for immunisation

PT of animals and monoclonal antibody development.

PS Claim 2; Page 1; 13pp; Japanese.

CC Probes comprising part of the sequence can be used to identify

CC Class I genes. The DNA can be expressed for immunisation of

CC animals and prodn. of monoclonal antibodies specific for the HLA-C

CC antigen. See also Q12116 (same patent) and J03112486 and J03112487.

CC Sequence 1101 BP; 215 A; 335 C; 379 G; 172 T;

SQ

alignment\_scores:

Quality: 44.00 Length: 10  
Ratio: 4.889 Gaps: 0  
Percent Similarity: 90.000 Percent Identity: 90.000

alignment\_block:

US-08-653-294-24 x Q12117

Align seg 1/1 to: Q12117 from: 1 to: 1101

1 ArgGluSerLeuArgAsnLeuArgGlyTyr 10  
|||||  
295 CGAGTGAGCGCTGCGGAACCTGCGCGGTAC 324

seq\_name: N\_Geneseq\_36:X22446

seq\_documentation\_block:

ID X22446 standard; DNA; 1323 BP.

AC X22446;

DT 20-MAY-1999 (first entry)

DE Human CPRM2 DNA.

KW CPRM1; human; chemoprotector regulated modifier; genotoxin; cytotoxin;

KW CPRM2; CPRM3; MCPRM1; allele s; allele r; gene therapy; antibody;

KW detoxification; detection; mutant; cancer; carcinogenic; aflatoxin;

KW treatment; neurodegeneration; Alzheimer's disease; malondialdehyde;

KW cyclo-oxygenase reaction; prostaglandin metabolism; ss.

OS Homo sapiens.

FT Key Location/Qualifiers

FT cds 74..994

```

FT FT WO9859055-A2.
PN PN
PD PD
PF PF
PR PR
PA PA
PI PI
PI Praml C, Schwab M;
DR WPI: 99-081282/07.
DR P-PSDB: W93087.
PT New chemoprotector regulated modifier proteins - useful in the
PT detoxification of metabolites, particularly aflatoxins
PS Claim 2; Fig 8; 31pp; German.
CC This sequence encodes a novel protein designated CPRM2 (chemoprotector
CC regulated modifier) that inhibits the effects of genotoxic and/or
CC cytotoxic substances. The invention describes the isolation of CPRM1,
CC CPRM2, CPRM3, mCPRM1, allele s, and mCPRM1, allele r. Transformants
CC containing such proteins are used to produce recombinant proteins and
CC nucleic acids for use in gene therapy and to raise antibodies which are
CC used to (i) detoxify geno- and/or cyto-toxic metabolites and (ii) detect
CC presence of CPRM proteins (or nucleic acid encoding them, including
CC detection of mutations indicative of increased risk of cancer, e.g.
CC prenatal). The proteins can be used to eliminate carcinogenic
CC aflatoxins and to treat neurodegeneration, e.g. Alzheimer's disease,
CC associated with production of malondialdehyde, a toxic product or the
CC cyclo-oxygenase reaction and prostaglandin metabolism.
SQ Sequence 1323 BP; 331 A; 368 C; 338 G; 286 T;

alignment_scores:
Quality: 39.00 Length: 10
Ratio: 4.333 Gaps: 0
Percent Similarity: 90.000 Percent Identity: 70.000

alignment_block:
US-08-653-294-24 x X22446/rev ..
Align seg 1/1 to reverse of: X22446 from: 1 to: 1323

1 ArgGluSerLeuArgAsnLeuArgGlyTyr 10
:::|||||:::||||| |||||||
1058 AAGAAGAGTGTGAGAACAAAGAGGTTAC 1029

seq_name: N_Geneseq_36:X22450

seq_documentation_block:
ID X22450 standard; DNA; 1324 BP.
AC X22450;
DT 20-MAY-1999 (first entry)
DE Human CPRM2 mutant DNA.
KW CPRM1; mouse; chemoprotector regulated modifier; genotoxin; cytotoxin;
KW CPRM2; CPRM3; mCPRM1; allele s; allele r; gene therapy; antibody;
KW detoxification; detection; mutant; cancer; carcinogenic; aflatoxin;
KW treatment; neurodegeneration; Alzheimer's disease; malondialdehyde;
KW cyclo-oxygenase reaction; prostaglandin metabolism; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT CDS
FT 74..223
FT /*tag- a
FT /product- "CPRM2 mutant"

PN PN
PD PD
PF PF
PR PR
PA (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
PI Praml C, Schwab M;
DR WPI: 99-081282/07.
DR P-PSDB: W93091.
PT New chemoprotector regulated modifier proteins - useful in the
PT detoxification of metabolites, particularly aflatoxins
PS Disclosure; Fig 12; 31pp; German.
CC This sequence encodes a novel protein designated CPRM2 which is a mutant

```

```

CC chemoprotector regulated modifier that inhibits the effects of
CC genotoxic and/or cytotoxic substances. The invention describes the
CC isolation of CPRM1, CPRM2, CPRM3, mCPRM1, allele s, and mCPRM1,
CC allele r. Transformants containing such proteins are used to produce
CC recombinant proteins and nucleic acids for use in gene therapy and to
CC raise antibodies which are used to (i) detoxify geno- and/or cyto-toxic
CC metabolites and (ii) detect presence of CPRM proteins (or nucleic acid
CC encoding them, including detection of mutations indicative of increased
CC risk of cancer, e.g. prenatal). The proteins can be used to eliminate
CC carcinogenic aflatoxins and to treat neurodegeneration, e.g. Alzheimer's
CC disease, associated with production of malondialdehyde, a toxic product
CC or the cyclo-oxygenase reaction and prostaglandin metabolism.
SQ Sequence 1324 BP; 332 A; 368 C; 338 G; 286 T;

alignment_scores:
Quality: 39.00 Length: 10
Ratio: 4.333 Gaps: 0
Percent Similarity: 90.000 Percent Identity: 70.000

alignment_block:
US-08-653-294-24 x X22450/rev ..
Align seg 1/1 to reverse of: X22450 from: 1 to: 1324

1 ArgGluSerLeuArgAsnLeuArgGlyTyr 10
:::|||||:::||||| |||||||
1059 AAGAAGAGTGTGAGAACAAAGAGGTTAC 1030

seq_name: N_Geneseq_36:T32217

seq_documentation_block:
ID T32217 standard; cDNA to mRNA; 1490 BP.
AC T32217;
DT 07-FEB-1997 (first entry)
DE Acyl-ACP-thioesterase coding sequence.
KW Acyl-ACP-thioesterase; AAT; cuphea leptopoda; aliphatic acid synthesis;
KW middle chain aliphatic acid; ss.
OS Cuphea leptopoda hemsl.
FH Key Location/Qualifiers
FT CDS
FT 30..1289
FT /*tag- a
FT /product- Acyl-ACP-thioesterase

PN PN
PD PD
PF PF
PR 22-DEC-1994; 320547.
PR 22-DEC-1994; JP-320547.
PA (NORO ) NORINSUISANSHO CHUGOKU NOGYO SHIKENBACH.
DR WPI: 96-365584/37.
DR P-PSDB: W02020.
PT DNA encoding Cuphea leptopoda Hemsl. acyl-ACP-thio:esterase - for
PT synthesis of middle chain aliphatic acids, useful in pharmaceutical
PT prepus. and foods
PT Claim 2; Page 6; 7pp; Japanese.
CC This sequence represents the coding sequence for the Cuphea leptopoda
CC acyl-ACP-thioesterase (AAT). This sequence was isolated by synthesizing
CC a AAT cDNA, and using that sequence to probe for a fragment of the AAT
CC gene (see T32217). The isolated fragment was then used to isolate the
CC AAT gene, which was then amplified. AAT is a key enzyme in the synthesis
CC of middle chain aliphatic acids, and as such is useful in pharmaceutical
CC preparations, and foods.
SQ Sequence 1490 BP; 351 A; 332 C; 402 G; 405 T;

alignment_scores:
Quality: 38.00 Length: 9
Ratio: 4.222 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 66.667

alignment_block:
US-08-653-294-24 x T32217/rev ..
Align seg 1/1 to reverse of: T32217 from: 1 to: 1490

```

```

2 GluSerLeuArgAsnLeuArgGlyTyr 10
:::||||:||||:||||:||||:||||:
1472 AATCAATCCGGAACATGCGAGGATAC 1446

seq_name: N_Geneseq_36:V35447

seq_documentation_block:
ID V35447 standard; CDNA; 1659 BP.
AC V35447;
DT 01-OCT-1998 (first entry)
DE Rat truncated rSK3 coding sequence.
KW Small conductance calcium-activated potassium channel protein 1;
KW rSK3; rat; potassium ion flux; ss.
OS Rattus sp.
PN WO9811139-A1.
PD 19-MAR-1998.
PF 10-SEP-1997; U16033.
PR 17-APR-1997; US-045233.
PR 11-SEP-1996; US-026451.
PR 07-MAR-1997; US-040052.
PA (ICAG-) ICAGEN INC.
PA (UYOR-) UNIV OREGON HEALTH SCI.
PI Adelman JP, Bond CT, Maylie J, Silvia CP;
DR WPI; 98-207332/18.
DR P-PSDB; W63703.
PT DNA encoding calcium-activated potassium channel - useful in assays
PT to identify compounds which increase or decrease potassium ion flux
PS Claim 3; Page 105; 151pp; English.
CC This sequence encodes the rat small conductance calcium-activated
CC potassium channel protein 3 (rSK3) of the invention. The proteins of the
CC invention are monomers of a calcium-activated potassium channel, where
CC the monomer: (i) has a calculated molecular weight of between 40 and
CC 80 kDa; and (ii) has a unit conductance of between 2 and 60 pS when the
CC monomer is in the functional polymeric form of a potassium chain and is
CC expressed in a Xenopus oocyte. Antibodies specific for the protein, and
CC probes specific for the DNA can be used to detect the presence of the
CC protein or DNA sequences in a sample. Host cells expression of the
CC protein can be used in assays to identify compounds which increase or
CC decrease the potassium ion flux through the protein. The transfect host
CC cell can also be used for the recombinant production of the protein. The
CC DNA sequences can also be used for determine mutations in the SK and IK
CC genes in a computer system. The proteins encoded by the SK and IK genes
CC can be used in a computer system for determining their three dimensional
CC structure, which is useful for determining ligands that bind to the
CC proteins.
SQ Sequence 1659 BP; 404 A; 499 C; 398 G; 358 T;

alignment_scores:
Quality: 38.00 Length: 10
Ratio: 3.800 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 70.000

alignment_block:
US-08-653-294-24 x V35447/rev ..
Align seg 1/1 to reverse of: V35447 from: 1 to: 1659

1 ArgGluSerLeuArgAsnLeuArgGlyTyr 10
||||:||||:||||:||||:||||:
62 CGAGAGCGCTGAGCGGCTGAGGGGCTTC 33

seq_name: N_Geneseq_36:V35458

seq_documentation_block:
ID V35458 standard; CDNA; 1674 BP.
AC V35458;
DT 01-OCT-1998 (first entry)
DE Truncated human hSK3 coding sequence.
KW Small conductance calcium-activated potassium channel protein 3;
KW hSK3; human; potassium ion flux; ss.
OS Homo sapiens.

alignment_scores:
Quality: 38.00 Length: 10
Ratio: 3.800 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 70.000

alignment_block:
US-08-653-294-24 x V35447/rev ..
Align seg 1/1 to reverse of: V35447 from: 1 to: 1659

1 ArgGluSerLeuArgAsnLeuArgGlyTyr 10
||||:||||:||||:||||:||||:
62 CGAGAGCGCTGAGCGGCTGAGGGGCTTC 33

seq_name: N_Geneseq_36:V35458

seq_documentation_block:
ID V35458 standard; CDNA; 1674 BP.
AC V35458;
DT 01-OCT-1998 (first entry)
DE Truncated human hSK3 coding sequence.
KW Small conductance calcium-activated potassium channel protein 3;
KW hSK3; human; potassium ion flux; ss.
OS Homo sapiens.

alignment_scores:
Quality: 38.00 Length: 10
Ratio: 3.800 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 70.000

alignment_block:
US-08-653-294-24 x V35458/rev ..
Align seg 1/1 to reverse of: V35458 from: 1 to: 1674

1 ArgGluSerLeuArgAsnLeuArgGlyTyr 10
||||:||||:||||:||||:||||:
62 CGGAGCGCGCTGAGCGGCTGAGGGGCTTC 33

seq_name: N_Geneseq_36:V35472

seq_documentation_block:
ID V35472 standard; CDNA; 2224 BP.
AC V35472;
DT 01-OCT-1998 (first entry)
DE Rat rSK3 coding sequence.
KW Small conductance calcium-activated potassium channel protein 3;
KW rSK3; rat; potassium ion flux; ss.
OS Rattus sp.
FH Key Location/Qualifiers
FT CDS 2..2200
FT /tag= a
PN WO9811139-A1.
PD 19-MAR-1998.
PF 10-SEP-1997; U16033.
PR 17-APR-1997; US-045233.
PR 11-SEP-1996; US-026451.
PR 07-MAR-1997; US-040052.
PA (ICAG-) ICAGEN INC.
PA (UYOR-) UNIV OREGON HEALTH SCI.
PI Adelman JP, Bond CT, Maylie J, Silvia CP;
DR WPI; 98-207332/18.
DR P-PSDB; W63715.
PT DNA encoding calcium-activated potassium channel - useful in assays

```

CC decrease the potassium ion flux through the protein. The transfected host  
CC cell can also be used for the recombinant production of the protein. The  
CC DNA sequences can also be used for determine mutations in the SK and IK  
CC genes in a computer system. The proteins encoded by the SK and IK genes  
CC can be used in a computer system for determining their three dimensional  
CC structure, which is useful for determining ligands that bind to the  
CC proteins.  
CC  
SQ Sequence 2452 BP; 552 A; 814 C; 617 G; 479 T;

alignment\_scores:  
Quality: 38.00 Length: 10  
Ratio: 3.800 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 70.000

alignment\_block:  
US-08-653-294-24 x V35473/rev ..

Align seq 1/1 to reverse of: V35473 from: 1 to: 2462

alignment_scores:		
Quality:	38.00	Length: 10
Ratio:	3.800	Gaps: 0
Percent Similarity:	100.000	Percent Identity: 70.000

```
alignment_block:
US-08-653-294-24 x V35473/rev
```

Align seq 1/1 to reverse of: v35473 from: 1 to: 2462

1 ArgGluSerLeuArgAsnLeuArgGlyTyr 10  
|||:::||||:::|||||:::  
50 CGGGAGCGCTGAGCGGCTGAGGGCTC 82

seq name: N Geneseq 36:X08414

seq\_documentation\_block:

ID X08414 standard; cDNA; 2521 BP.

DT	28-JUN-1999 (first entry)
DE	Human small conductance calcium activated potassium channel gene.
DW	Human small conductance calcium activated potassium channel;
KW	hKCA3/KCNH3; bipolar disease; schizophrenia; treatment; diagnosis;
KW	detection; transgenic animal; gene therapy;
KW	neuropsychiatric disorder; neurological disorder;
KW	neuromuscular disorder; immunological disorder; ss.

home sapiens,	Location/Qualifiers
key	287..2482
CDS	/tag= a
	/product= Small con
	potassium
	2482..2487
polyA_signal	/tag= b
	/label= Poly-A sig
polyA_site	2509..2521
	/tag= c
	/label= Poly-A req

WO9903889-A1.

28-JAN-1999.

14-JUL-1998; U14902.

08-JAN-1998; US-070741.

15-JUL-1997; US-052556.

( REGC ) UNIV CALIFORNIA.

Chandy KG, Fantino E, Ga

WPI; 99-132165/11.

P-PSDB; W96312.

New human small conducta

useful for diagnosis, tr

schizophrenia and bipolar

Claim 3; Figure 5; 95pp;  
Disclosures associated with

Disorders associated with

conductance calcium-activated  
comprise neuropsychiatric

comprise neuropsychiatric immunological disorders.

Immunological Disorders;  
schizophrenia. Transgeni

as a transgene in their

as models for studying the

Antibodies (Ab) directed

or conservative variants

those at risk from these

**such disorders. Use of t**

these disorders is also





```

FT FT /tag= k
FT FT /standard_name= "ORF K11"
FT FT /product= "protein of unknown function"
FT FT /note= "homologous to the FixU gene"
FT CDS complement (434107..434433)
FT FT /tag= l
FT FT /standard_name= "ORF K12"
FT FT /product= "protein of unknown function"
FT CDS complement (434517..434711)
FT FT /tag= m
FT FT /standard_name= "ORF K13"
FT FT /product= "ferrodoxin/ferrodoxin-like protein"
FT FT /note= "homologous to the FdxN gene"
FT CDS complement (434753..436234)
FT FT /tag= n
FT FT /standard_name= "ORF K14"
FT FT /gene= "nifB"
FT FT /product= "protein involved in FeMo co-factor biosynthesis"
FT CDS complement (436460..438130)
FT FT /tag= o
FT FT /standard_name= "ORF K15"
FT FT /gene= "nifA"
FT FT /product= "positive regulator of nif, fix and other genes"
FT CDS complement (438297..438590)
FT FT /tag= p
FT FT /standard_name= "ORF K16"
FT FT /gene= "fixX"
FT FT /product= "protein required for nitrogenase activity"
FT CDS complement (438605..439912)
FT FT /tag= q
FT FT /standard_name= "ORF K17"
FT FT /gene= "fixC"
FT FT /product= "protein required for nitrogenase activity"
FT CDS complement (439923..441032)
FT FT /tag= r
FT FT /standard_name= "ORF K18"
FT FT /gene= "fixB"
FT FT /product= "protein required for nitrogenase activity"
FT CDS complement (441042..441899)
FT FT /tag= s
FT FT /standard_name= "ORF K19"
FT FT /gene= "fixA"
FT FT /product= "protein required for nitrogenase activity"
FT CDS complement (442316..442636)
FT FT /tag= t
FT FT /standard_name= "ORF K20"
FT FT /product= "protein of unknown function"
FT CDS complement (443313..443879)
FT FT /tag= u
FT FT /standard_name= "ORF K21"
FT FT /product= "protein of unknown function"
FT CDS complement (444337..445029)
FT FT /tag= v
FT FT /standard_name= "ORF K22"
FT FT /product= "ferrodoxin-like protein"
FT FT /note= "homologous to the NifQ gene"
FT CDS complement (445088..446602)
FT FT /tag= w
FT FT /standard_name= "ORF K23"
FT FT /gene= "dctA"
FT FT /product= "C4-dicarboxylate transport protein"
FT FT /note= "homologous to the DctAI gene"
FT CDS complement (446599..447843)
FT FT /tag= x
FT FT /standard_name= "ORF L1"
FT FT /product= "cytochrome P450-like protein"
FT FT /note= "homologous to the CamC gene"
FT CDS complement (447844..448500)
FT FT /tag= y
FT FT /standard_name= "ORF L2"
FT FT /product= "gamma-hexachlorocyclohexane-dechlorinase-like

```

```

FT FT /note= "homologous to the LinA gene"
FT CDS complement (448497..450203)
FT FT /tag= z
FT FT /standard_name= "ORF L3"
FT FT /product= "putative protein with degradative function"
FT CDS complement (450341..451396)
FT FT /tag= aa
FT FT /standard_name= "ORF L4"
FT FT /product= "luciferase alpha-subunit-like protein"
FT FT /note= "homologous to the LuxA gene"
FT CDS complement (452980..454494)
FT FT /tag= ab
FT FT /standard_name= "ORF L6"
FT FT /gene= "nifD"
FT FT /product= "alpha-subunit of FeMo protein of nitrogenase"
FT CDS complement (454590..456131)
FT FT /tag= ac
FT FT /standard_name= "ORF L7"
FT FT /gene= "nifK"
FT FT /product= "beta-subunit of FeMo protein of nitrogenase"
FT CDS complement (456187..457677)
FT FT /tag= ad
FT FT /standard_name= "ORF L8"
FT FT /product= "protein involved in FeMo co-factor biosynthesis"
FT CDS complement (457687..459096)
FT FT /tag= ae
FT FT /standard_name= "ORF L9"
FT FT /product= "protein involved in FeMo co-factor biosynthesis"
FT CDS complement (459093..459575)
FT FT /tag= af
FT FT /standard_name= "ORF L10"
FT FT /product= "protein of unknown function"
FT FT /note= "homologous to the NifX gene"
FT CDS complement (459579..460067)
FT FT /tag= ag
FT FT /standard_name= "ORF L11"
FT FT /product= "protein of unknown function"
FT CDS complement (460501..460920)
FT FT /tag= ah
FT FT /standard_name= "ORF L12"
FT FT /product= "protein similar to part of the Fe protein of nitrogenase"
FT CDS complement (461228..461545)
FT FT /tag= ai
FT FT /standard_name= "ORF L13"
FT FT /product= "protein of unknown function"
FT CDS complement (463201..464739)
FT FT /tag= aj
FT FT /standard_name= "ORF L14"
FT FT /product= "peptidase-like protein"
FT FT /note= "homologous to the BI-MPP gene"
FT CDS complement (464736..466079)
FT FT /tag= ak
FT FT /standard_name= "ORF L15"
FT FT /product= "processing protease-like protein"

```

## alignment\_scores:

```

Quality: 37.00 Length: 9
Ratio: 4.625 Gaps: 0
Percent Similarity: 88.889 Percent Identity: 88.889

```

## alignment\_block:

```

US-08-653-294-24 x v30458_0

```

```

Align seg 1/1 to: v30458_0 from: 1 to: 110000

```

```

1 ArgGluSerLeuArgAsnLeuArgGly 9

```

15658 CGCGAAAGTCTTAGGAATCTCTTGGG 15684  
|||||

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run On: February 8, 2000, 04:05:42 ; Search time 133.56 Seconds  
(without alignments)  
1.773 Million cell updates/sec

Title: US-08-653-294-25  
Perfect score: 50  
Sequence: 1 RENLRTALRY 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 188963 seqs, 23686106 residues

Total number of hits satisfying chosen parameters: 188963

Minimum DB seq length: 0  
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database : A\_Geneseq\_36.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50	100.0	10	R83095	HLA-B2702 CTL modu
2	50	100.0	10	R95426	Peptide B2702.75-84(T)
3	50	100.0	10	W33788	HLA-B2702.75-8
4	50	100.0	20	R92809	HLA-B2702 CTL modu
5	50	100.0	20	R92910	HLA-B2702 CTL modu
6	50	100.0	20	W33792	Peptide B2702.84-7
7	50	100.0	20	W33793	Peptide B2702.84-7
8	44	88.0	10	R41208	Peptide fragment o
9	44	88.0	10	R83062	HLA-B2702 CTL modu
10	44	88.0	10	R95413	Alphal-helix of HL
11	44	88.0	10	W07512	T-cell modulating
12	44	88.0	10	W47265	Immunomodulatory p
13	44	88.0	10	W33784	Peptide B2702.75-8
14	44	88.0	15	R92912	HLA-B2702 CTL modu
15	44	88.0	15	W33795	Peptide B2702.70-8
16	44	88.0	20	R92907	HLA-B2702 CTL modu
17	44	88.0	20	R92908	HLA-B2702 CTL modu
18	44	88.0	20	R95428	HLA-B2702.84-75-84
19	44	88.0	20	W33778	Immunomodulating d
20	44	88.0	20	W33791	Peptide B2702.84-7
21	44	88.0	25	R41205	Peptide fragment o
22	44	88.0	25	R48286	Peptide fragment o
23	44	88.0	25	R83090	HLA-B2702 CTL modu
24	44	88.0	25	R83093	HLAB38 CTL modulat
25	44	88.0	25	R95416	HLA-B2702.60-84. C
26	44	88.0	25	R95422	HLAB38.6084. Comps
27	44	88.0	25	W33794	Peptide B2702.60-8
28	44	88.0	184	Y06801	Peptide Seq ID No:
29	44	88.0	362	R03142	Sequence of HLA-Bw
30	44	88.0	362	R03144	Sequence of HLA-B5
31	44	88.0	362	R12463	HLA-Bw53 exon. HLA
32	43	86.0	10	W07522	T-cell modulating
33	40	80.0	10	R41212	Peptide fragment o
34	40	80.0	10	R83075	HLA-B2702 CTL modu

35 40 80.0 10 1 R95423 HLA-B2705.75-84. C  
36 40 80.0 10 1 W33785 Peptide B2705.75-8  
37 40 80.0 17 1 R71442 Human HLA-B27-(62-  
38 40 80.0 17 1 R71443 Human [Phe74]-HLA-  
39 40 80.0 25 1 R41221 Peptide fragment o  
40 40 80.0 25 1 R83091 HLA-B2702 CTL modu  
41 40 80.0 25 1 R95417 HLA-B2705.60-84. C  
42 40 80.0 337 1 P70590 Sequence of the hu  
43 40 80.0 362 1 P70155 Sequence encoded b  
44 39 78.0 10 1 R83094 HLA-B2702 CTL modu  
45 39 78.0 10 1 R83096 HLA-B2702 CTL modu

#### ALIGNMENTS

RESULT 1  
R83095 ID R83095 standard; peptide; 10 AA.  
AC R83095;  
DT 16-MAY-1996 (first entry)  
DE HLA-B2702 CTL modulating peptide (B2702.75-84(T)).  
KW Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;  
KW immunosuppressant; graft versus host disorder; transplantation; therapy;  
KW class I MHC; HLA-B2702.  
OS Synthetic.  
PN W09526979-A1.  
PD 12-OCT-1995.  
PF 05-APR-1995; U04349.  
PR 05-APR-1994; US-222851.  
PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
PI Clayberger C, Krensky AM, Parham P;  
DR WPI; 95-358582/46.  
PT Extension of acceptance period of transplants from MHC unmatched  
PT donor hosts - using Class I B7-5/84 MHC antigen of the recipient  
PT host  
PS Example 14; Page 34; 80pp; English.  
CC R83061-R83085, R83090-R83096 and R92907-R92913 represent fragments of  
CC class I major histocompatibility complex (MHC) antigens. This sequence  
CC corresponds to residues 75-84 of the alpha-1 domain of the class I MHC  
CC HLA-B2702. This sequence showed no inhibitory effect upon cytotoxic T  
CC lymphocytes (CTLs). These sequences can be used to extend the period of  
CC acceptance by a recipient of a transplant from an MHC unmatched donor.  
CC The peptides are administered to a patient in conjunction with  
CC a subtherapeutic amount of an immunosuppressant. This is administered to  
CC the patient for a limited period of time (compared to the lifetime  
CC administration for current treatments). The peptides particularly  
CC modulate (or inhibit) the activity of the CTLs of the patient.  
SQ Sequence 10 AA;

Query Match 100.0%; Score 50; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.00033;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RENLRTALRY 10  
| | | | | | | | | |  
DB 1 RENLRTALRY 10

RESULT 2  
R95426 ID R95426 standard; peptide; 10 AA.  
AC R95426;  
DT 12-NOV-1996 (first entry)  
DE HLA-B2702.75-84(T).  
KW HLA; p74; alpha1-helix; human-leucocyte-associated antigen; inhibitor;  
KW T-cell lysate; membrane protein; mammal; heat shock protein; Hsc70; APC;  
KW B cell; calcium influx; cytotoxic T lymphocyte; CTL; differentiation;  
KW cytolysis; antigen presenting cell.  
OS Synthetic.  
PN Key  
FT misc\_difference 6 Location/Qualifiers

/note= "I6T mutation"

FT W09513288-A1.  
 PN 18-MAY-1995.  
 PF 10-NOV-1994; U12985.  
 PR 10-NOV-1993; US-150493.  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI Clayberger C, Krensky AM;  
 DR WPI: 95-194027/25.  
 PT Compens. comprising lymphoid surface membrane proteins - which may  
 PT inhibit cytolytic activity and differentiation of CTLs.  
 PS Example: Page 11: 29pp; English.  
 CC R95413, and R95413-R95431 represent palindromes and fragments of  
 CC human-leucocyte-associated antigens. This sequence represents the  
 CC HLA-B2702.75-84(T). These sequences can be used to isolate the protein  
 CC p74 from a T-cell lysate. p74 is a T-cell surface membrane protein  
 CC associated with T-cell activation in mammalian T-cells, and is also  
 CC immunologically cross reactive with the heat shock protein Hsc70. p74 is  
 CC found in a limited number of cell types, but is particularly expressed on  
 CC B and T cells. p74 can be isolated by lysis of a suitable cell with an  
 CC amphoteric detergent, and then passed through an affinity column  
 CC containing a covalently bound HLA-B2702 palindromic peptide.  
 CC Compositions comprising the extracellular fragment of p74 combined with  
 CC HLA-B2702.80-84 (see R95416), induces calcium influx, and inhibits  
 CC cytotoxic T lymphocyte (CTL) differentiation or cytotoxicity. Candidate  
 CC compounds can be screened for their effect on the cytolytic activity of  
 CC T-cells, by combining them with the extracellular portion of p74 and  
 CC determining the amount of binding between the candidate compound and p74.  
 CC Modulation of CTL activity can be inhibited in a cellular composition  
 CC containing T-cells and antigen presenting cells (APCs), by adding to the  
 CC mix the extracellular portion of p74, in an amount sufficient to compete  
 CC with p74 for the binding of the p74 ligand.  
 SQ Sequence 10 AA;

Query Match 100.0%; Score 50; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.00033;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RENLRTALRY 10  
 |||||  
 Db 1 RENLRTALRY 10

RESULT 3  
 W33788  
 ID W33788 standard; peptide; 10 AA.  
 AC W33788:  
 DT 19-JUN-1998 (first entry)  
 DE Peptide B2702.75-84T80 tested for immunomodulating activity.  
 KW Immunomodulating dimer; immunosuppressant drug; CTL activation;  
 KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;  
 KW rejection.  
 OS Synthetic.  
 OS Homo sapiens.  
 PN W09744351-A1.  
 PD 27-NOV-1997.  
 PF 22-MAY-1997: U08689.  
 PR 24-MAY-1996: US-653294.  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI Beulow R, Clayberger C, Krensky AM;  
 DR WPI: 98-086530/08.  
 PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B  
 PT alpha-1 domain, used for preventing rejection of transplants or  
 PT treating autoimmune diseases

PS Example 1: Page 19: 41pp; English.  
 CC Peptides W33784-98 and W33778-9 were assayed for their immunomodulating  
 CC activity. A peptide-type compound or variant is claimed which has  
 CC immunomodulating activity, including the N-terminal acylated and/or  
 CC C-terminal amidated or esterified forms of up to 60 amino acids, where  
 CC the peptide-type compound comprises the formula: A-B, where A, B =  
 CC (R aa76-77L) (aa79-84) or (aa84-79) (1aa77-76R); aa76 = E or V; aa77 =  
 CC D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a hydrophobic or  
 CC small amino acid; aa82 = R or L; aa83 = G or R; and aa represents amino

CC acid. The sequence in the brackets may optionally be absent or truncated  
 CC at any peptide type bond within the brackets. The compounds comprise  
 CC amino acid sequences related to a Class I HLA-B alpha domain (positions  
 CC 79-84). They can be used to inhibit cytotoxic T-lymphocytes (CTL) from  
 CC undesirably attacking cells in a host or in vitro. They can also be  
 CC used in combination with antigenic peptides or proteins of interest to  
 CC activate CTLs. They can also inhibit the proliferation of T cells in  
 CC response to anti-CD3. The peptide can be used for preventing rejection  
 CC of transplants or for treating autoimmune diseases, e.g. diabetes,  
 CC rheumatoid arthritis and lupus erythematosus. The products can also be  
 CC used for detection and diagnosis.  
 SQ Sequence 10 AA;

Query Match 100.0%; Score 50; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.00033;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RENLRTALRY 10  
 |||||  
 Db 1 RENLRTALRY 10

RESULT 4  
 R92909  
 ID R92909 standard; peptide; 20 AA.  
 AC R92909:  
 DT 16-MAY-1996 (first entry)  
 DE HLA-B2702 CTL modulating peptide (B2702.84-75/75-84(T)).  
 KW Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;  
 KW immunosuppressant; graft versus host disorder; transplantation; therapy;  
 KW class I MHC; HLA-B2702.  
 OS Synthetic.  
 PN W09526579-A1.  
 PD 12-OCT-1995.  
 PF 05-APR-1995; U04349.  
 PR 05-APR-1994; US-222851.  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI Clayberger C, Krensky AM, Parham P;  
 DR WPI: 95-358582/46.  
 PT Extension of acceptance period of transplants from MHC unmatched  
 PT donor hosts - using Class I B75-84 MHC antigen of the recipient  
 PT host

PS Example 15: Page 36: 80pp; English.  
 CC R83061-R83085, R83090-R83096 and R92907-R92913 represent fragments of  
 CC class I major histocompatibility complex (MHC) antigens. This sequence  
 CC is an inverted dimer of residues 75-84 of the alpha-1 domain of the class  
 CC I MHC HLA-B2702. These sequences can be used to extend the period of  
 CC acceptance by a recipient of a transplant from an MHC unmatched donor.  
 CC The peptides are administered to a patient in conjunction with a  
 CC subtherapeutic amount of an immunosuppressant. This is administered to  
 CC the patient for a limited period of time (compared to the lifetime  
 CC administration for current treatments). The peptides particularly  
 CC modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)  
 CC of the patient.  
 SQ Sequence 20 AA;

Query Match 100.0%; Score 50; DB 1; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 0.0007;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RENLRTALRY 10  
 |||||  
 Db 11 RENLRTALRY 20

RESULT 5  
 R92910  
 ID R92910 standard; peptide; 20 AA.  
 AC R92910:  
 DT 16-MAY-1996 (first entry)  
 DE HLA-B2702 CTL modulating peptide (B2702.84-75(T)/75-84(T)).

KW Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;  
 KW immunosuppressant; graft versus host disorder; transplantation; therapy;  
 KW class I MHC; HLA-B\*2702.  
 OS Synthetic.  
 PN W03528979-A1.  
 PD 12-OCT-1995.  
 PF 05-APR-1995; U04349.  
 PR 05-APR-1994; US-222851.  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI Clayberger C, Krensky AM, Parham P;  
 DR WPI; 95-358562/46.  
 PT Extension of acceptance period of transplants from MHC unmatched  
 PT donor hosts - using Class I B\*5-84 MHC antigen of the recipient  
 PT host  
 PS Example 15; Page 36; 80pp; English.  
 CC R83061-R83085, R83090-R83096 and R92907-R92913 represent fragments of  
 CC class I major histocompatibility complex (MHC) antigens. This sequence  
 CC is an inverted dimer of residues 75-84 of the alpha-1 domain of the class  
 CC I MHC HLA-B\*2702. These sequences can be used to extend the period of  
 CC acceptance by a recipient of a transplant from an MHC unmatched donor.  
 CC The peptides are administered to a patient in conjunction with a  
 CC subtherapeutic amount of an immunosuppressant. This is administered to  
 CC the patient for a limited period of time (compared to the lifetime  
 CC administration for current treatments). The peptides particularly  
 CC modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)  
 CC of the patient.  
 CC Sequence 20 AA;

Query Match 100.0%; Score 50; DB 1; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 0.0007;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RENLRTALRY 10  
 Db 11 RENLRTALRY 20

## RESULT 6

W33792  
 ID W33792 standard; peptide; 20 AA.  
 AC W33792;  
 DT 19-JUN-1998 (first entry)  
 DE Peptide B2702.84-75/75-84 tested for immunomodulating activity.  
 KW Immunomodulating dimer; immunosuppressant drug; CTL activation;  
 KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;  
 KW rejection.  
 OS Synthetic.  
 OS Homo sapiens.  
 PN W09744351-A1.  
 PD 27-NOV-1997.  
 PF 22-MAY-1997; U08689.  
 PR 24-MAY-1996; US-653294.  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI Beulow R, Clayberger C, Krensky AM;  
 DR WPI; 98-086530/08.  
 PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B  
 PT alpha-1 domain, used for preventing rejection of transplants or  
 PT treating autoimmune diseases  
 PS Example 1; Page 19; 41pp; English.  
 CC Peptides W33784-98 and W33778-9 were assayed for their immunomodulating  
 CC activity. A peptide-type compound or variant is claimed which has  
 CC immunomodulating activity, including the N-terminal acylated and/or  
 CC C-terminal amidated or esterified forms of up to 60 amino acids, where  
 CC the peptide-type compound comprises the formula: A-B, where A, B =  
 CC (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or V; aa77 =  
 CC D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a hydrophobic or  
 CC small amino acid; aa82 = R or L; aa83 = G or R; and aa represents amino  
 CC acid. The sequence in the brackets may optionally be absent or truncated  
 CC at any peptide type bond within the brackets. The compounds comprise  
 CC amino acid sequences related to a Class I HLA-B alpha domain (positions  
 CC 79-84). They can be used to inhibit cytotoxic T-lymphocytes (CTL) from  
 CC undesirably attacking cells in a host or in vitro. They can also be  
 CC used in combination with antigenic peptides or proteins of interest to  
 CC activate CTLs. They can also inhibit the proliferation of T cells in  
 CC response to anti-CD3. The peptide can be used for preventing rejection  
 CC of transplants or for treating autoimmune diseases, e.g. diabetes,  
 CC rheumatoid arthritis and lupus erythematosus. The products can also be  
 CC used for detection and diagnosis.

CC used in combination with antigenic peptides or proteins of interest to  
 CC activate CTLs. They can also inhibit the proliferation of T cells in  
 CC response to anti-CD3. The peptide can be used for preventing rejection  
 CC of transplants or for treating autoimmune diseases, e.g. diabetes,  
 CC rheumatoid arthritis and lupus erythematosus. The products can also be  
 CC used for detection and diagnosis.  
 SQ Sequence 20 AA;

Query Match 100.0%; Score 50; DB 1; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 0.0007;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RENLRTALRY 10  
 Db 11 RENLRTALRY 20

## RESULT 7

W33793  
 ID W33793 standard; peptide; 20 AA.  
 AC W33793;  
 DT 19-JUN-1998 (first entry)  
 DE Peptide B2702.84-75/75-84 tested for immunomodulating activity.  
 KW Immunomodulating dimer; immunosuppressant drug; CTL activation;  
 KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;  
 KW rejection.  
 OS Synthetic.  
 OS Homo sapiens.  
 PN W09744351-A1.  
 PD 27-NOV-1997.  
 PF 22-MAY-1997; U08689.  
 PR 24-MAY-1996; US-653294.  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI Beulow R, Clayberger C, Krensky AM;  
 DR WPI; 98-086530/08.  
 PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B  
 PT alpha-1 domain, used for preventing rejection of transplants or  
 PT treating autoimmune diseases  
 PS Example 1; Page 19; 41pp; English.  
 CC Peptides W33784-98 and W33778-9 were assayed for their immunomodulating  
 CC activity. A peptide-type compound or variant is claimed which has  
 CC immunomodulating activity, including the N-terminal acylated and/or  
 CC C-terminal amidated or esterified forms of up to 60 amino acids, where  
 CC the peptide-type compound comprises the formula: A-B, where A, B =  
 CC (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or V; aa77 =  
 CC D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a hydrophobic or  
 CC small amino acid; aa82 = R or L; aa83 = G or R; and aa represents amino  
 CC acid. The sequence in the brackets may optionally be absent or truncated  
 CC at any peptide type bond within the brackets. The compounds comprise  
 CC amino acid sequences related to a Class I HLA-B alpha domain (positions  
 CC 79-84). They can be used to inhibit cytotoxic T-lymphocytes (CTL) from  
 CC undesirably attacking cells in a host or in vitro. They can also be  
 CC used in combination with antigenic peptides or proteins of interest to  
 CC activate CTLs. They can also inhibit the proliferation of T cells in  
 CC response to anti-CD3. The peptide can be used for preventing rejection  
 CC of transplants or for treating autoimmune diseases, e.g. diabetes,  
 CC rheumatoid arthritis and lupus erythematosus. The products can also be  
 CC used for detection and diagnosis.

Query Match 100.0%; Score 50; DB 1; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 0.0007;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RENLRTALRY 10  
 Db 11 RENLRTALRY 20

## RESULT 8

R41208

ID R41208 standard; peptide; 10 AA.  
AC R41208;  
DT 15-MAR-1994 (first entry)  
DE Peptide fragment of Class I HLA peptide.  
KW Human leukocyte antigen; HLA; peptide; transplantation; neoplasia;  
KW parasitic disease; cytotoxic T lymphocyte; modulation.  
OS Synthetic.  
PN W09317699-A.  
PD 16-SEP-1993.  
PF 25-FEB-1993; U01758.  
PR 02-MAR-1992; US-844716.  
PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
PI Clayberger CA, Krensky AM;  
PT WPI: 93-303134/38.  
PT New peptide(s) based on Class I HLA antigen domains - used for  
PT modulating cytotoxic T-lymphocyte activity towards targets  
PS Claim 11; Page 54; 61pp; English.  
CC The peptide is used to modulate cytotoxic T-lymphocyte (CTL)  
CC activity, either by inhibition or stimulation. It can be used  
CC for inhibiting CTL toxicity in transplantations, for inducing CTL  
CC activity in parasitic diseases and neoplasia and in studies on viral  
CC infection. The peptide can also be used for identifying CTLs which  
CC bind to it and removing subsets of CTLs from a T-cell composition.  
CC This peptide sequence is more commonly found within larger peptide  
CC compounds of not more than 30 amino acids in length.  
SQ Sequence 10 AA;

Query Match 88.0%; Score 44; DB 1; Length 10;  
Best Local Similarity 90.0%; Pred. No. 0.0051;  
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RENLRLALRY 10  
DB 1 RENLRLALRY 10  
||||| |||||

RESULT 9  
ID R83062 standard; peptide; 10 AA.  
AC R83062;  
DT 16-MAY-1996 (first entry)  
DE HLA-B2702 CTL modulating peptide (B2702.75-84).  
KW Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;  
KW immunosuppressant; graft versus host disorder; transplantation; therapy;  
KW Class I MHC; HLA-B2702.  
OS Synthetic.  
PN W09526979-A1.  
PD 12-OCT-1995.  
PF 05-APR-1995; U04349.  
PR 05-APR-1994; US-222851.  
PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
PI Clayberger C, Krensky AM, Parham P;  
DR WPI: 95-358582/46.  
PT Extension of acceptance period of transplants from MHC unmatched  
PT donor hosts - using Class I B\*5-84 MHC antigen of the recipient  
PT host  
PS Claim 15; Page 9; 80pp; English.  
CC R83061-R83085, R83090-R83096 and R92907-R92913 represent fragments of  
CC Class I major histocompatibility complex (MHC) antigens. This sequence  
CC corresponds to residues 75-84 of the alpha-1 domain of the class I MHC  
CC HLA-B2702. These sequences can be used to extend the period of  
CC acceptance by a recipient of a transplant from an MHC unmatched donor.  
CC The peptides are administered to a patient in conjunction with a  
CC subtherapeutic amount of an immunosuppressant. This is administered to  
CC the patient for a limited period of time (compared to the lifetime  
CC administration for current treatments). The peptides particularly  
CC modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)  
CC of the patient.  
SQ Sequence 10 AA;

Query Match 88.0%; Score 44; DB 1; Length 10;  
Best Local Similarity 90.0%; Pred. No. 0.0051;  
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RENLRLALRY 10  
DB 1 RENLRLALRY 10  
||||| |||||

RESULT 11  
ID W07512 standard; peptide; 10 AA.  
AC W07512;  
DT 04-AUG-1997 (first entry)  
DE T-cell modulating peptide #1.  
KW T-cell modulator; autoimmune disease; tissue destruction; alaph-domain;  
KW mammal; major histocompatibility complex; MHC class I; antigen; perforin;  
KW insulin-dependent diabetes mellitus; multiple sclerosis; inflammation;  
KW rheumatoid arthritis; psoriasis; pemphigus vulgaris; Sjogren's disease;  
KW thyroid disease; Hashimoto's thyroiditis; myasthenia gravis; granzyme;  
KW autologous target cell; cytokine release; T cell activation; therapy.

Query Match 88.0%; Score 44; DB 1; Length 10;  
Best Local Similarity 90.0%; Pred. No. 0.0051;  
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RENLRLALRY 10  
DB 1 RENLRLALRY 10  
||||| |||||

RESULT 10  
R95413  
ID R95413 standard; peptide; 10 AA.  
AC R95413;  
DT 12-NOV-1996 (first entry)  
DE Alaph-helix of HLA-B2702.  
KW HLA; p74; alaph-helix; human-leucocyte-associated antigen; inhibitor;  
KW T-cell lysate; membrane protein; mammal; heat shock protein; Hsc70; APC;  
KW B cell; calcium influx; cytotoxic T lymphocyte; CTL; differentiation;  
KW cytolysis; antigen presenting cell.  
OS Synthetic.  
PN W09513288-A1.  
PD 18-MAY-1995.  
PF 10-NOV-1994; U12985.  
PR 10-NOV-1993; US-150493.  
PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
PI Clayberger C, Krensky AM;  
DR WPI: 95-194027/25.  
PT Compsns. comprising lymphoid surface membrane proteins - which may  
PT inhibit cytolytic activity and differentiation of CTLs.  
PS Example: Page 11; 29pp; English.  
CC This sequence represents the alaph-helix of the  
CC human-leucocyte-associated antigen B2702 (HLA-B2702). This sequence,  
CC epitopes, and palindromes of it (such as R95428) can be used to isolate  
CC the protein p74 from a T-cell lysate. p74 is a T-cell surface membrane  
CC protein associated with T-cell activation in mammalian T-cells, and is  
CC also immunologically cross reactive with the heat shock protein Hsc70.  
CC p74 is found in a limited number of cell types, but is particularly  
CC expressed on B and T cells. p74 can be isolated through an affinity  
CC cell with an amphoteric detergent, and then passed through an affinity  
CC column containing a covalently bound HLA-B2702 palindromic peptide.  
CC Compositions comprising the extracellular fragment of p74 combined with  
CC HLA-B2702.60-84 (see R95416), induces calcium influx, and inhibits  
CC cytotoxic T lymphocyte (CTL) differentiation or cytolysis. Candidate  
CC compounds can be screened for their effect on the cytolytic activity of  
CC T-cells, by combining them with the extracellular portion of p74 and  
CC determining the amount of binding between the candidate compound and p74.  
CC Modulation of CTL activity can be inhibited in a cellular composition  
CC containing T-cells and antigen presenting cells (APCs), by adding to the  
CC mix the extracellular portion of p74, in an amount sufficient to compete  
CC with p74 for the binding of the p74 ligand.  
SQ Sequence 10 AA;



OS Synthetic.  
 PN WO9635443-A1.  
 PD 14-NOV-1996.  
 PF 05-APR-1996; U04710.  
 PR 12-MAY-1995; US-440504.  
 PA (SANG-) SANGSTAT MEDICAL CORP.  
 PI Buelow R;  
 DR WPI: 96-518410/51.  
 PT Treatment of auto-immune disease by admin. of peptide(s) corresp. to  
 PT major histocompatibility complex antigens - esp. for delaying onset  
 PT of clinical symptoms of insulin dependent diabetes by modulating T  
 PT cell mediated attack on target cells  
 PS Claim 7; Page 20; 24pp; English.  
 CC WO7512-W07518 represent T-cell modulating peptides that can be used in  
 CC the method of the invention. These sequences are based on a portion of  
 CC the generic peptide corresponding to residues 70-91 of the alpha-domain  
 CC of the major histocompatibility complex (MHC) class I antigen (see  
 CC WO7510). The method is for affecting the course of an autoimmune disease  
 CC involving T-cell mediated destruction of tissue in mammals. These  
 CC peptides are used especially to treat insulin-dependent diabetes  
 CC mellitus, preferably being administered during the pre-clinical stage to  
 CC delay onset of the disease. Other diseases that can be treated are  
 CC multiple sclerosis, rheumatoid arthritis, psoriasis, pemphigus vulgaris,  
 CC Sjogren's disease, thyroid disease, Hashimoto's thyroiditis, myasthenia  
 CC gravis, etc. The peptides modulate T-cell mediated attack on autologous  
 CC target cells, and may also reduce inflammation, swelling, and release of  
 CC cytokines, perforins, granzymes etc. associated with T cell activation.  
 CC Sequence 10 AA;  
 SQ

Query Match 88.0%; Score 44; DB 1; Length 10;  
 Best Local Similarity 90.0%; Pred. No. 0.0051;  
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RENLRLALRY 10  
 Db 1 RENLRLALRY 10

RESULT 12  
 W47265  
 ID W47265 standard; peptide; 10 AA.  
 AC W47265;  
 DT 22-MAY-1998 (first entry)  
 DE Immunomodulatory peptide.  
 KW Immunomodulator; Class I HLA-B alpha-1 domain; inhibition;  
 KW transplant rejection; treatment; autoimmune disease.  
 OS Homo sapiens.  
 OS Synthetic.  
 FS Key Location/Qualifiers  
 FT Misc\_difference 1..10 /note= "at least one of the amino acids is the  
 FT D-isomer  
 FT  
 PN WO9744052-A1.  
 PD 27-NOV-1997.  
 PF 23-APR-1997; U06705.  
 PR 22-MAY-1996; US-651650.  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI Clayberger C, Krensky AM;  
 DR WPI: 98-018220/02.  
 PT Novel immunomodulatory peptide-type compound - useful for inhibiting  
 PT transplant rejection  
 PS Claim 10; Page 36; 41pp; English.  
 CC The present sequence is an immunomodulatory peptide, which  
 CC comprises a Class I HLA-B alpha-1 domain sequence. It can be used  
 CC in a pharmaceutical composition together with a subtherapeutic dose  
 CC of an immunosuppressant, to extend the period of acceptance of a  
 CC transplant from a major histocompatibility complex (MHC) unmatched  
 CC donor, i.e. to inhibit transplant rejection. It can also be used in  
 CC the treatment of autoimmune diseases.  
 CC Peptides using the D-form amino acids are more effective  
 CC immunomodulators than their diastereomers or enantiomers.  
 CC Sequence 10 AA;  
 SQ

Query Match 88.0%; Score 44; DB 1; Length 10;  
 Best Local Similarity 90.0%; Pred. No. 0.0051;  
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RENLRLALRY 10  
 Db 1 RENLRLALRY 10

RESULT 14  
 R92912  
 ID R92912 standard; peptide; 15 AA.  
 AC R92912;  
 DT 16-MAY-1996 (first entry)  
 DE HLA-B2702 CTL modulating peptide (B2702.70-84).  
 KW Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;  
 KW immunosuppressant; graft versus host disorder; transplantation; therapy;  
 KW class I MHC; HLA-B2702.

Query Match 88.0%; Score 44; DB 1; Length 10;  
 Best Local Similarity 90.0%; Pred. No. 0.0051;  
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RENLRLALRY 10  
 Db 1 RENLRLALRY 10

RESULT 13  
 W33784  
 ID W33784 standard; peptide; 10 AA.  
 AC W33784;  
 DT 19-JUN-1998 (first entry)  
 DE Peptide B2702.75-84 tested for immunomodulating activity.  
 KW immunomodulating dimer; immunosuppressant drug; CTL activation;  
 KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;  
 KW rejection.  
 OS Synthetic.  
 OS Homo sapiens.  
 PN WO9744351-A1.  
 PD 27-NOV-1997.  
 PF 22-MAY-1997; U08689.  
 PR 24-MAY-1996; US-653294.  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI Buelow R, Clayberger C, Krensky AM;  
 DR WPI: 98-086530/08.  
 PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B  
 PT alpha-1 domain, used for preventing rejection of transplants or  
 PT treating autoimmune diseases  
 PS Example 1; Page 19; 41pp; English.  
 CC Peptides W33784-98 and W33778-9 were assayed for their immunomodulating  
 CC activity. A peptide-type compound or variant is claimed which has  
 CC immunomodulating activity, including the N-terminal acylated and/or  
 CC C-terminal amidated or esterified forms of up to 60 amino acids, where  
 CC the peptide-type compound comprises the formula: A-B, where A, B =  
 CC (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or V; aa77 =  
 CC D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a hydrophobic or  
 CC small amino acid; aa82 = R or L; aa83 = G or R; and aa represents amino  
 CC acid. The sequence in the brackets may optionally be absent or truncated  
 CC at any peptide type bond within the brackets. The compounds comprise  
 CC amino acid sequences related to a Class I HLA-B alpha-1 domain (positions  
 CC 79-84). They can be used to inhibit cytotoxic T-lymphocytes (CTL) from  
 CC undesirably attacking cells in a host or in vitro. They can also be  
 CC used in combination with antigenic peptides or proteins of interest to  
 CC activate CTLs. They can also inhibit the proliferation of T cells in  
 CC response to anti-CD3. The peptide can be used for preventing rejection  
 CC of transplants or for treating autoimmune diseases, e.g. diabetes,  
 CC rheumatoid arthritis and lupus erythematosus. The products can also be  
 CC used for detection and diagnosis.  
 CC Sequence 10 AA;  
 SQ

OS Synthetic.  
 PN W09526979-A1.  
 PD 12-OCT-1995.  
 PF 05-APR-1995; U04349.  
 PR 05-APR-1994; US-222851.  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI Clayberger C, Krensky AM, Parham P;  
 DR WPI: 95-358582/46.  
 PT Extension of acceptance period of transplants from MHC unmatched  
 PT donor hosts - using Class I B75-84 MHC antigen of the recipient  
 PT host  
 PS Example 15; Page 36; 80pp; English.  
 CC R3061-R3085, R83090-R83096 and R92907-R92913 represent fragments of  
 CC class I major histocompatibility complex (MHC) antigens. This sequence  
 CC corresponds to residues 70-84 of the alpha-1 domain of the class I MHC  
 CC HLA-B\*2702. These sequences can be used to extend the period of  
 CC acceptance by a recipient of a transplant from an MHC unmatched donor.  
 CC The peptides are administered to a patient in conjunction with a  
 CC subtherapeutic amount of an immunosuppressant. This is administered to  
 CC the patient for a limited period of time (compared to the lifetime  
 CC administration for current treatments). The peptides particularly  
 CC modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)  
 CC of the patient.  
 SQ Sequence 15 AA;

Query Match 88.0%; Score 44; DB 1; Length 15;  
 Best Local Similarity 90.0%; Pred. No. 0.008;  
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RENLRALRY 10  
 ||||| ||||  
 Db 6 RENLRALRY 15

RESULT 15  
 W33795  
 ID W33795 standard; peptide; 15 AA.  
 AC W33795;  
 DT 19-JUN-1998 (first entry)  
 DE Peptide B2702.70-84 tested for immunomodulating activity.  
 KW Immunomodulating dimer; Immunosuppressant drug; CTL activation;  
 KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;  
 KW rejection.  
 OS Synthetic.  
 OS Homo sapiens.  
 PN W09744351-A1.  
 PD 27-NOV-1997.  
 PF 22-MAY-1997; U08689.  
 PR 24-MAY-1996; US-653294.  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI Beulow R, Clayberger C, Krensky AM;  
 DR WPI: 98-086530/08.  
 PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B  
 PT alpha-1 domain, used for preventing rejection of transplants or  
 PT treating autoimmune diseases  
 PS Example 1; Page 19; 41pp; English.  
 CC Peptides W33784-98 and W33778-9 were assayed for their immunomodulating  
 CC activity. A peptide-type compound or variant is claimed which has  
 CC immunomodulating activity, including the N-terminal acylated and/or  
 CC C-terminal amidated or esterified forms of up to 60 amino acids, where  
 CC the peptide-type compound comprises the formula: A-B, where A, B =  
 CC (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or V; aa77 =  
 CC D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a hydrophobic or  
 CC small amino acid; aa82 = R or L; aa83 = G or R; and aa represents amino  
 CC acid. The sequence in the brackets may optionally be absent or truncated  
 CC at any peptide type bond within the brackets. The compounds comprise  
 CC amino acid sequences related to a Class I HLA-B alpha1 domain (positions  
 CC 79-84). They can be used to inhibit cytotoxic T-lymphocytes (CTL) from  
 CC undesirably attacking cells in a host or in vitro. They can also be  
 CC used in combination with antigenic peptides or proteins of interest to  
 CC activate CTLs. They can also inhibit the proliferation of T cells in  
 CC response to anti-CD3. The peptide can be used for preventing rejection

CC of transplants or for treating autoimmune diseases, e.g. diabetes,  
 CC rheumatoid arthritis and lupus erythematosus. The products can also be  
 CC used for detection and diagnosis.  
 SQ Sequence 15 AA;

Query Match 88.0%; Score 44; DB 1; Length 15;  
 Best Local Similarity 90.0%; Pred. No. 0.008;  
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RENLRALRY 10  
 ||||| ||||  
 Db 6 RENLRALRY 15

Search completed: February 8, 2000, 04:05:43  
 Job time: 9360 sec

OM of: US-08-653-294-24 to: EST.\* out\_format : pfs

Date: Feb 8, 2000 6:23 AM

About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp  
-O=cpnl\_1/USPTO.spool/US08653294/runat\_04022000\_160700\_15770/app\_query.fasta.2  
-DB=EST -FMT=fastap -SUFFIX=rst -GAPOP=12.000 -GAPEXT=4.000  
-MINMATCH=0.100 -LOOPCL=0.000 -LOPEXT=0.000 -GAPOP=4.500  
-GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -GAPOP=6.000  
-FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000  
-DELEXT=7.000 -START=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DCCALIGN=200 -THR SCORE=pct -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pfs -NORM=ext -MINLEN=0 -MAXLEN=1000000 -USER=US08653294  
-NCPU=6 -ICPU=3 -NO\_XLPHY -WAIT -THREADS=1

Search information block:

Query: US-08-653-294-24

Query length: 10

Database: EST.\*

Database sequences: 4538634

Database length: 1887831982

Search time (sec): 7600.090000

score\_list:

Sequence	Strd	Orig	zScore	Escore	Len	Documentation
gb_est12:AA319533	+	51.00	134.48	0.0457	103	AA319533 EST121772 Adrenal gland
gb_est13:AA361477	+	51.00	188.68	0.0961	189	AA361477 EST170761 T-cell lympho
gb_est13:AA352603	+	51.00	187.31	0.1145	218	AA352603 EST160074 Activated T-c
gb_est12:AA294911	+	51.00	185.66	0.1415	259	AA294911 EST100074 Pancreas tum
gb_est13:AA352960	+	51.00	185.26	0.1489	270	AA352960 EST161101 Activated T-c
gb_est8:AA058454	+	51.00	182.17	0.2213	373	AA058454 z167407.r1 Stratagene
gb_est11:AA224068	+	51.00	181.58	0.2389	397	AA224068 z111402.r1 Stratagene
gb_est10:AA160317	+	51.00	181.39	0.2448	405	AA160317 z056c07.r1 Stratagene
gb_est7:W40489	+	51.00	180.88	0.2613	427	W40489 zc84b01.r1 Pancreatic is
gb_est12:AA310808	+	51.00	180.53	0.2733	443	AA310808 EST181593 Jurkat T-cell
gb_est9:Cl0310	+	51.00	179.86	0.2977	475	Cl0310 Cl0310 Human placenta cl
gb_est17:AA663896	+	51.00	179.78	0.3008	479	AA663896 ae74d01.s1 Stratagene
gb_est35:AA039796	+	51.00	179.31	0.3194	503	AA039796 DKF2p434B1912.r1 434
gb_est11:AA263135	+	51.00	176.02	0.4875	710	AA263135 PMY0598 KG1-a Lambda Z
gb_est27:AA242623	-	44.00	161.23	3.25	209	AA242623 tf34g01.x1 NCI_CGAP_B
gb_est23:AA097437	-	44.00	160.03	3.79	237	AA097437 qb61e08.x1 NCI_CGAP_B
gb_est14:AA449607	+	44.00	156.00	6.33	360	AA449607 zx06h02.r1 Soares tota
gb_est1:TS5088	+	44.00	156.00	6.35	361	TS5088 yb45c09.r1 Stratagene fe
gb_est11:AA227277	+	44.00	155.64	6.65	375	AA227277 zr20h03.r1 Stratagene
gb_est31:AA1706797	+	44.00	154.81	7.40	409	AA1706797 UI-R-AA1-aab-h-09-0-UI
gb_est36:AA1706862	+	44.00	154.67	7.53	415	AA1706862 UI-R-AA1-aac-c-04-0-UI
gb_est23:AA089801	-	44.00	154.22	7.98	435	AA089801 qa11a09.x1 NCI_CGAP_B
gb_est11:AA243206	-	44.00	154.13	8.07	439	AA243206 zr26c11.r1 Stratagene
gb_est26:AA1356710	-	44.00	153.69	8.55	460	AA1356710 qy22d12.x1 NCI_CGAP_B
gb_est15:AA477506	+	44.00	153.30	8.75	469	AA477506 zu35f05.s1 Soares ovary
gb_est4:H23327	+	44.00	153.30	8.98	479	H23327 ym52e08.r1 Soares infant
gb_est26:AA849373	+	44.00	153.18	9.12	485	AA849373 EST192140 Normalized r
gb_est36:AA924062	+	44.00	152.85	9.51	502	AA924062 UI-R-A1-du-e-07-0-UI.s
gb_est37:AA1939506	-	44.00	152.72	9.68	509	AA1939506 qy22d12.x5 NCI_CGAP_B
gb_est26:AA123583	+	44.00	152.30	10.22	532	AA123583 EST229271 Normalized r
gb_est36:AA945159	+	44.00	152.01	10.59	548	AA945159 EST200658 Normalized r
gb_est11:AA1232284	+	44.00	150.41	13.01	648	AA1232284 EST228972 Normalized r
gb_est10:AA0248102	-	42.00	144.36	28.27	553	AA0248102 HS 2015.A1.H09 MR CIT
gb_est27:AA140530	+	41.00	146.52	21.43	297	AA140530 EST238823 Normalized r
gb_est29:AA1602357	+	41.00	145.34	25.56	343	AA1602357 UI-R-AA0-wn-h-05-0-UI.
gb_est39:AA100680	+	41.00	143.23	32.68	419	AA100680 z190502.r1 Stratagene
gb_est38:AA060987	+	41.00	141.73	39.60	490	AA060987 UI-M-BH1-ann-c-01-0-UI
gb_est9:AA083156	+	40.00	141.70	39.76	331	AA083156 zn08a04.r1 Stratagene
gb_est9:AA121088	+	40.00	139.63	51.85	411	AA121088 zm22a05.r1 Stratagene
gb_gss11:AA0333000	-	40.00	138.21	62.24	477	AA0333000 HS-5003.A2.A08_r7 RPL1
gb_est10:AA158109	-	40.00	137.85	65.13	495	AA158109 zo57h10.s1 Stratagene
gb_est10:AA160714	+	40.00	137.22	70.66	529	AA160714 zo77e11.r1 Stratagene

gb\_gss13:AA0468916 - 40.00 137.02 72.47 540 ! AA0468916 HS-5140.B2.H09.T7A  
gb\_est11:AA203261 - 40.00 132.13 135.60 900 ! AA203261 zx55d11.r1 Soares f  
gb\_est4:F14616 + 39.00 143.90 29.97 177 ! F14616 SSOAH02 Porcine small

seq\_name: gb\_est12:AA319533

seq\_documentation\_block:

LOCUS AA319533 103 bp mRNA EST 19-APR-1997

DEFINITION EST121772 Adrenal gland tumor Homo sapiens cDNA 5' end similar to  
similar to major histocompatibility complex, class I, B  
(GB:W16102), mRNA sequence.

ACCESSION AA319533

VERSION AA319533

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 103)

REFERENCE

AUTHORS

Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,  
Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,  
White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,  
Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,  
Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S.,  
Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr.,  
Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,  
Moreno-Palauques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,  
Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,  
Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,F.,  
Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,  
Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,  
He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,  
Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Weissner,P.S., Olsen,H.,  
Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M.,  
Dillon,P.J., Fannon,M.R., Rosen,C.A., Haseitine,W.A., Fields,C.,  
Fraser,C.M. and Venter,J.C.

TITLE Initial assessment of human gene diversity and expression patterns

based upon 83 million nucleotides of cDNA sequence

Nature 377 (6547 Suppl), 3-174 (1995)

JOURNAL

MEDLINE

COMMENT

On Sep 12, 1996 this sequence version replaced gi:1393672.

Other\_ESTS: TH0169519

Contact: Kerlavage, AR

Bioinformatics

The Institute for Genomic Research

9712 Medical Center Drive, Rockville, MD 20850 USA

Tel: 3018699056

Fax: 3018699423

Email: arkerlavet@tigr.org

For clone availability, additional sequence and expression

information related to this EST, please check the TIGR Human Gene

Index (<http://www.tigr.org/cdb/hgi/hgi.html>)

Seq primer: M13 Reverse

Location/Qualifiers

source 1..103

/organism="Homo sapiens"

/db\_xref="ATCC (inhost):120083"

/db\_xref="taxon:9606"

/clone\_lib="adrenal gland tumor"

/db\_stag="adult"

/note="Organ: adrenal gland; Vector: pBluescript SK-;

Site\_1: EcoRI; Site\_2: XhoI"

BASE COUNT 26 a 35 c 29 g 13 t

ORIGIN

alignment\_scores:

Quality: 51.00 Length: 10

Ratio: 5.100 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-08-653-294-24 x AA319533 ..

Align seg 1/1 to: AA319533 from: 1 to: 103

1 ArgGluSerLeuArgAsnLeuArgGlyTyr 10  
|||||  
27 CGAGAGACCTCGCGAACCTTCGGCGCTAC 56

seq\_name: gb\_est13:AA361477

seq\_documentation\_block: 189 bp mRNA EST 21-APR-1997  
LOCUS AA361477 T-cell lymphoma Homo sapiens cDNA 5' end similar to  
DEFINITION EST70761 T-cell lymphoma Homo sapiens cDNA 5' end similar to  
similar to major histocompatibility complex, class I, Bw62.3, mRNA  
sequence.

ACCESSION AA361477  
VERSION AA361477.1 GI:2013795  
KEYWORDS EST.  
SOURCE human.

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 189)

REFERENCE  
AUTHORS Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A., Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man,Wai,C., Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D., Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S., Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr., Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M., Moreno-Palancas,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M., Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R., Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y., Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J., Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A., He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K., Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Weissner,P.S., Olsen,H., Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M., Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and Venter,J.C.  
Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence

Nature 377 (6547 suppl), 3-174 (1995)  
12140200  
On Sep 12, 1996 this sequence version replaced gi:1404737.  
Other ESTs: THCL69519

JOURNAL MEDLINE COMMENT  
The Institute for Genomic Research  
9712 Medical Center Drive, Rockville, MD 20850 USA  
Tel: 3018699056  
Fax: 3018699423  
Email: arkerlavet@tigr.org  
For clone availability, additional sequence and expression information related to this EST, please check the TIGR Human Gene Index (<http://www.tigr.org/tdb/hgi/hgi.html>)  
Seq primer: M13 Reverse.

FEATURES  
Source  
1..189  
/organism="Homo sapiens"  
/db\_xref="ATCC (inhost):165623"  
/db\_xref="taxon:9606"  
/clone\_lib="T-cell lymphoma"  
/cell\_type="T-lymphocyte"  
/note="Vector: pBluescript SK-; Site\_1: EcoRI; Site\_2: XhoI"  
BASE COUNT 45 a 60 c 55 g 25 t 4 others

alignment\_scores:  
Quality: 51.00 Length: 10  
Ratio: 5.100 Gaps: 0  
Percent similarity: 100.000 Percent identity: 100.000

alignment\_block:  
US-08-653-294-24 x AA361477

Align seg 1/1 to: AA361477 from: 1 to: 189

1 ArgGluSerLeuArgAsnLeuArgGlyTyr 10  
|||||  
37 CGAGAGACCTCGCGAACCTTCGGCGNTAC 66

seq\_name: gb\_est13:AA352603

seq\_documentation\_block: 218 bp mRNA EST 21-APR-1997  
LOCUS AA352603  
DEFINITION EST60621 Activated T-cells XX Homo sapiens cDNA 5' end similar to  
similar to major histocompatibility complex, class I, Bw62.3, mRNA  
sequence.

ACCESSION AA352603  
VERSION AA352603.1 GI:2004923  
KEYWORDS EST.  
SOURCE human.

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 218)

REFERENCE  
AUTHORS Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A., Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man,Wai,C., Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D., Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S., Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr., Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M., Moreno-Palancas,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M., Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R., Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y., Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J., Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A., He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K., Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Weissner,P.S., Olsen,H., Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M., Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and Venter,J.C.  
Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence

Nature 377 (6547 suppl), 3-174 (1995)  
12140200  
On May 8, 1995 this sequence version replaced gi:800964.  
Other ESTs: THCL72938

JOURNAL MEDLINE COMMENT  
The Institute for Genomic Research  
9712 Medical Center Drive, Rockville, MD 20850 USA  
Tel: 3018699056  
Fax: 3018699423  
Email: arkerlavet@tigr.org  
For clone availability, additional sequence and expression information related to this EST, please check the TIGR Human Gene Index (<http://www.tigr.org/tdb/hgi/hgi.html>)  
Seq primer: M13 Reverse.

FEATURES  
source  
Location/Qualifiers  
1..218  
/organism="Homo sapiens"  
/db\_xref="ATCC (inhost):152802"  
/db\_xref="taxon:9606"  
/clone\_lib="Activated T-cells XX"  
/cell\_type="T-lymphocyte"  
/dev\_stage="adult"  
/note="Vector: pBluescript SK-; Site\_1: EcoRI; Site\_2: XhoI"  
BASE COUNT 52 a 70 c 62 g 28 t 6 others

```

alignment_scores:
  Quality: 51.00      Length: 10
  Ratio: 5.100      Gaps: 0
  Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
  US-08-653-294-24 x AA352603
  ..
  Align seg 1/1 to: AA352603 from: 1 to: 218

1 ArgGluSerLeuArgAsnLeuArgGlyTyr 10
|||||
39 CGAGAGAGCCTGCGGAACTGCGGCGCTAC 68

seq_name: gb_est12:AA294911

seq_documentation_block:
  LOCUS AA294911 259 bp mRNA EST 18-APR-1997
  DEFINITION EST100074 Pancreas tumor I Homo sapiens cDNA 5' end similar to
  similar to major histocompatibility complex, class I, Bw62.3, mRNA
  sequence.
  ACCESSION AA294911
  VERSION AA294911.1 GI:1947266
  KEYWORDS EST.
  SOURCE human.
  ORGANISM Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
  Eutheria; Primates; Catarrhini; Hominidae; Homo.
  REFERENCE 1 (bases 1 to 259)
  AUTHORS Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
  Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
  White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,
  Clifton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
  Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S.,
  Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr.,
  Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
  Moreno-Palauques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
  Phillips,C.A., Ryder,S., Scott,J.L., Saudek,D.M., Shirley,R.,
  Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
  Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
  Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,
  He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
  Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Weissner,P.S., Olsen,H.,
  Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M.,
  Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
  Fraser,C.M. and Venter,J.C.
  TITLE Initial assessment of human gene diversity and expression patterns
  based upon 83 million nucleotides of cDNA sequence
  JOURNAL Nature 377 (6547 Suppl), 3-174 (1995)
  MEDLINE 12140200
  COMMENT On Nov 29, 1993 this sequence version replaced gi:430148.
  Other_ESTs: THC172938
  Contact: Kerlavage, AR
  Bioinformatics
  The Institute for Genomic Research
  9712 Medical Center Drive, Rockville, MD 20850 USA
  Tel: 3018699056
  Fax: 3018699423
  Email: arkerlavet@tigr.org
  For clone availability, additional sequence and expression
  information related to this EST, please check the TIGR Human Gene
  Index (http://www.tigr.org/tdb/hgi/hgi.html)
  Seq primer: M13 Reverse.
  Location/Qualifiers
  1..259
  /organism="Homo sapiens"
  /db_xref="ATCC (inhost):190413"
  /db_xref="taxon:9606"
  /clone_lib="Pancreas tumor I"
  /dev_stage="adult"
  /note="Organ: pancreas; Vector: pBluescript SK-; Site_1:
  EcoRI; Site_2: XhoI"

BASE COUNT 60 a 87 c 30 t 7 others

ORIGIN

alignment_scores:
  Quality: 51.00      Length: 10
  Ratio: 5.100      Gaps: 0
  Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
  US-08-653-294-24 x AA294911
  ..
  Align seg 1/1 to: AA294911 from: 1 to: 259

1 ArgGluSerLeuArgAsnLeuArgGlyTyr 10
|||||
42 CGAGAGAGCCTGCGGAACTGCGGCGGENTAC 71

seq_name: gb_est13:AA352960

seq_documentation_block:
  LOCUS AA352960 270 bp mRNA EST 21-APR-1997
  DEFINITION EST61101 Activated T-cells XX Homo sapiens cDNA 5' end similar to
  similar to major histocompatibility complex, class I, B61, mRNA
  sequence.
  ACCESSION AA352960
  VERSION AA352960.1 GI:2005353
  KEYWORDS EST.
  SOURCE human.
  ORGANISM Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
  Eutheria; Primates; Catarrhini; Hominidae; Homo.
  REFERENCE 1 (bases 1 to 270)
  AUTHORS Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
  Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
  White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,
  Clifton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
  Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S.,
  Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr.,
  Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
  Moreno-Palauques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
  Phillips,C.A., Ryder,S., Scott,J.L., Saudek,D.M., Shirley,R.,
  Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
  Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
  Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,
  He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
  Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Weissner,P.S., Olsen,H.,
  Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M.,
  Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
  Fraser,C.M. and Venter,J.C.
  TITLE Initial assessment of human gene diversity and expression patterns
  based upon 83 million nucleotides of cDNA sequence
  JOURNAL Nature 377 (6547 Suppl), 3-174 (1995)
  MEDLINE 12140200
  COMMENT On May 8, 1995 this sequence version replaced gi:801269.
  Other_ESTs: THC172938
  Contact: Kerlavage, AR
  Bioinformatics
  The Institute for Genomic Research
  9712 Medical Center Drive, Rockville, MD 20850 USA
  Tel: 3018699056
  Fax: 3018699423
  Email: arkerlavet@tigr.org
  For clone availability, additional sequence and expression
  information related to this EST, please check the TIGR Human Gene
  Index (http://www.tigr.org/tdb/hgi/hgi.html)
  Seq primer: M13 Reverse.
  Location/Qualifiers
  1..270
  /organism="Homo sapiens"
  /db_xref="ATCC (inhost):153240"
  /db_xref="taxon:9606"
  /clone_lib="Activated T-cells XX"
  /cell_type="T-lymphocyte"
  FEATURES
  source
  /organism="Homo sapiens"
  /db_xref="ATCC (inhost):153240"
  /db_xref="taxon:9606"
  /clone_lib="Activated T-cells XX"
  /cell_type="T-lymphocyte"

```

/dev\_stage="adult"  
/note="Vector: pBluescript SK-; Site\_1: EcoRI; Site\_2:  
XhoI"

BASE COUNT 62 a 80 c 88 g 35 t 5 others  
ORIGIN

alignment\_scores:  
Quality: 51.00 Length: 10  
Ratio: 5.100 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-08-653-294-24 x AA352960

Align seg 1/1 to: AA352960 from: 1 to: 270

1 ArgGlusLeuArgAsnLeuArgGlyTyr 10

|||||  
182 CGAGAGAGCTCGGAACCTCGCGGCTAC 211

seq\_name: gb\_est8:AA058454

seq\_documentation\_block:

LOCUS AA058454 373 bp mRNA EST 01-DEC-1996  
DEFINITION z167d07.r1 Stratagene colon (#937204) Homo sapiens cDNA clone  
IMAGE:509677 5' similar to gb:M24039.cdsl HLA CLASS I  
HISTOCOMPATIBILITY ANTIGEN, B-18 B\*1801 ALPHA (HUMAN);, mRNA  
sequence.

ACCESSION AA058454

VERSION AA058454.1 GI:1551280

KEYWORDS EST.

SOURCE human.

ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B.,  
Chissoe, S., Dietrich, N., DuBuque, T., Favello, A., Gish, W.,  
Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N.,  
Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L.,  
Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J.,  
Trevisan, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R.  
and Marra, M.

TITLE Generation and analysis of 280,000 human expressed sequence tags

JOURNAL Genome Res. 6 (9), 807-828 (1996)

MEDLINE 97044478

COMMENT

On Sep 12, 1996 this sequence version replaced gi:1393357.

Contact: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Insert Length: 1592 Std Error: 0.00

Seq primer: -28M13 rev2 from Amersham

High quality sequence stop: 268.

Location/Qualifiers

1..373

/organism="Homo sapiens"

/db\_xref="GDB:3813009"

/db\_xref="taxon:9606"

/clone="IMAGE:509677"

/clone\_lib="Stratagene colon (#937204)"

/lab\_host="SOLR cells (kanamycin resistant)"

/note="Organ: colon; Vector: pBluescript SK-; Site\_1:

EcoRI; Site\_2: XhoI; Cloned unidirectionally. Primer:

Oligo dT. 1-84 colonic epithelial cell line. Average

insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor

sequence: 5' GAATTCGGCAGAG 3' -3' adaptor sequence: 5'

CTCGAGTTTTTTTTTTTTTTT 3'

BASE COUNT 73 a 119 c 116 g 57 t 8 others  
ORIGIN

alignment\_scores:

Quality: 51.00 Length: 10  
Ratio: 5.100 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-08-653-294-24 x AA058454

Align seg 1/1 to: AA058454 from: 1 to: 373

1 ArgGlusLeuArgAsnLeuArgGlyTyr 10

|||||  
229 CGAGAGAGCTCGGAACCTCGCGGCTAC 258

seq\_name: gb\_est11:AA224068

seq\_documentation\_block:

LOCUS AA224068 397 bp mRNA EST 19-FEB-1997  
DEFINITION zrlf02.r1 Stratagene hnt neuron (#937233) Homo sapiens cDNA clone  
IMAGE:648507 5' similar to gb:L22649.cdsl HLA CLASS I  
HISTOCOMPATIBILITY ANTIGEN, B-39 B\*3901 ALPHA (HUMAN);, mRNA  
sequence.

ACCESSION AA224068

VERSION AA224068.1 GI:1844610

KEYWORDS EST.

SOURCE human.

ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 397)

AUTHORS

Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B.,  
Chissoe, S., Dietrich, N., DuBuque, T., Favello, A., Gish, W.,  
Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N.,  
Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L.,  
Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J.,  
Trevisan, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R.  
and Marra, M.

TITLE Generation and analysis of 280,000 human expressed sequence tags

JOURNAL Genome Res. 6 (9), 807-828 (1996)

MEDLINE 97044478

COMMENT

On Nov 29, 1993 this sequence version replaced gi:430429.

Contact: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Seq primer: -28M13 rev1 ET from Amersham

High quality sequence stop: 295.

Location/Qualifiers

1..397

/organism="Homo sapiens"

/db\_xref="GDB:558245"

/db\_xref="taxon:9606"

/clone="IMAGE:648507"

/clone\_lib="Stratagene hnt neuron (#937233)"

/lab\_host="SOLR (kanamycin resistant)"

/note="Vector: pBluescript SK-; Site\_1: EcoRI; Site\_2:

XhoI; Cloned unidirectionally. Primer: Oligo dT.

Differentiated, post mitotic hnt neurons. Average insert

size: 1.5 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5'

GAATTCGGCAGAG 3' -3' adaptor sequence: 5'

CTCGAGTTTTTTTTTTTTTTT 3'

77 a 129 c 131 g 56 t 4 others

BASE COUNT

ORIGIN

## alignment\_scores:

Quality: 51.00 Length: 10  
Ratio: 5.100 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-08-653-294-24 x AA224068 ..

Align seg 1/1 to: AA224068 from: 1 to: 397

1 ArgGluSerLeuArgAsnLeuArgGlyTyr 10  
|||||  
297 CGAGAGCGCTCGNAACCTGCGCGCTAC 326

seq\_name: gb\_est10:AA160317

## seq\_documentation\_block:

LOCUS AA160317 405 bp mRNA EST 09-MAR-1998  
DEFINITION z056c07.r1 Stragatene pancreas (#937208) Homo sapiens cDNA clone IMAGE:590892 5' similar to gb:U11265 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-35 B\*3501 ALPHA (HUMAN);, mRNA sequence.

ACCESSION AA160317.1 GI:1734956

VERSION AA160317

KEYWORDS EST.

SOURCE human.

## ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
Eutheria; Primates; Catarrhini; Hominiidae; Homo.

## REFERENCE

1 (bases 1 to 405)  
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S., Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wyllie, T., Waterston, R. and Wilson, R.  
WashU-NCI human EST Project  
Unpublished (1997)

## TITLE

On Sep 12, 1996 this sequence version replaced gi:1406818.

## JOURNAL

Contact: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

Insert Length: 1671 Std Error: 0.00

Seq primer: -28M13 rev2 from Amersham

High quality sequence stop: 332.

## FEATURES

source

1..405

/organism="Homo sapiens"

/db\_xref="GDB:4622278"

/db\_xref="taxon:9606"

/clone="IMAGE:590892"

/lab\_host="Stragatene pancreas (#937208)"

/note="Organ: pancreas (kanamycin resistant)"

ECORI; Site:2; XhoI; Cloned unidirectionally. Primer:

Oligo dT. Pancreatic adenocarcinoma cell line. Average

insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor

sequence: 5' GAATTCGCGACGAG 3' -3' adaptor sequence: 5'

CTCGAGTTTTTTTTTTTTTTT 3'

BASE COUNT 80 a 132 c 132 g 58 t 3 others

ORIGIN

## alignment\_scores:

Quality: 51.00 Length: 10  
Ratio: 5.100 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-08-653-294-24 x AA160317 ..

Align seg 1/1 to: AA160317 from: 1 to: 405

1 ArgGluSerLeuArgAsnLeuArgGlyTyr 10  
|||||  
281 CGAGAGCGCTCGGAACCTGCGCGCTAC 310

seq\_name: gb\_est7:W40489

## seq\_documentation\_block:

LOCUS W40489 427 bp mRNA EST 20-MAY-1996  
DEFINITION z084b01.r1 Pancreatic Islet Homo sapiens cDNA clone IMAGE:328969 5' similar to gb:U11265 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-35 B\*3501 ALPHA (HUMAN);, mRNA sequence.

ACCESSION W40489

VERSION W40489.1 GI:1324496

KEYWORDS EST.

SOURCE human.

## ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
Eutheria; Primates; Catarrhini; Hominiidae; Homo.

## REFERENCE

1 (bases 1 to 427)

Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B.,

Chisoe, S., Dietrich, N., DuBuque, T., Favello, A., Gish, W.,

Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, N.,

Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L.,

Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J.,

Trevaskis, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R.,

and Marra, M.

Generation and analysis of 280,000 human expressed sequence tags

Genome Res. 6 (9), 807-828 (1996)

97044478

On Jan 25, 1995 this sequence version replaced gi:637865.

Contact: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Seq primer: mob.REGA+ET

High quality sequence stop: 397.

FEATURES

source

1..427

/organism="Homo sapiens"

/db\_xref="GDB:1263173"

/db\_xref="taxon:9606"

/clone="IMAGE:328969"

/lab\_host="Pancreatic Islet"

/tissue\_type="pancreatic islet"

/note="Organ: pancreas (kanamycin resistant)"

ECORI; Site:2; XhoI; Cloned unidirectionally. Primer:

Takeda et al. Cloned unidirectionally. Primer: Oligo dT.

-5' adaptor sequence: 5' GAATTCGCGACGAG 3' -3' adaptor

sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'

BASE COUNT 87 a 141 c 134 g 62 t 3 others

ORIGIN

## alignment\_scores:

Quality: 51.00 Length: 10  
Ratio: 5.100 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-08-653-294-24 x W40489 ..

Align seg 1/1 to: W40489 from: 1 to: 427

1 ArgGluSerLeuArgAsnLeuArgGlyTyr 10

|||||

279 CGAGAGAGCTCGGAACCTCGCGGCTAC 308

seq\_name: gb\_est12:AA310808

seq\_documentation\_block: 443 bp mRNA EST 19-APR-1997  
 LOCUS AA310808  
 DEFINITION EST181593 Jurkat T-cells V Homo sapiens cDNA 5' end similar to similar to major histocompatibility complex, class I, B0704, mRNA sequence.  
 ACCESSION AA310808  
 VERSION AA310808.1 GI:1963136  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 1 (bases 1 to 443)  
 AUTHORS Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A., Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C., Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D., Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghegan,N.S., Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr., Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M., Moreno-Palancas,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M., Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R., Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y., Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J., Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A., He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K., Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Weissner,P.S., Olsen,H., Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M., Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and Venter,J.C.  
 TITLE Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence  
 JOURNAL Nature 377 (6547 Suppl), 3-174 (1995)  
 MEDLINE 12140200  
 COMMENT On Sep 12, 1996 this sequence version replaced gi:1397854.  
 Other\_ESTs: THC180721  
 Contact: Kerlavage, AR  
 Bioinformatics  
 The Institute for Genomic Research  
 9712 Medical Center Drive, Rockville, MD 20850 USA  
 Tel: 3018699056  
 Fax: 3018699423  
 Email: arkerlavetigr.org  
 For clone availability, additional sequence and expression information related to this EST, please check the TIGR Human Gene Index (<http://www.tigr.org/tdb/hgi/hgi.html>)  
 Seq primer: M13 Reverse

FEATURES  
 Location/Qualifiers  
 1..443

/organism="Homo sapiens"  
 /db\_xref="ATCC (inhost):156811"  
 /db\_xref="taxon:9606"  
 /clone\_lib="Jurkat T-cells V"  
 /cell\_type="T-lymphocyte"  
 /note="Vector: pBluescript SK-; Site\_1: EcoRI; Site\_2: XhoI"

BASE COUNT 86 a 144 c 143 g 64 t 6 others  
 ORIGIN  
 alignment\_scores:  
 Quality: 51.00 Length: 10  
 Ratio: 5.100 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
 US-08-653-294-24 x AA310808

Align seg 1/1 to: AA310808 from: 1 to: 443

1 ArgGluSerLeuArgAsnLeuArgGlyTyr 10  
 |||||  
 292 CGAGAGAGCTCGGAACCTCGCGGCTAC 321

seq\_name: gb\_est9:C18310

seq\_documentation\_block: 475 bp mRNA EST 02-OCT-1996  
 LOCUS C18310  
 DEFINITION C18310 Human placenta cDNA (TFujiwara) Homo sapiens cDNA clone GEN-560D07 5', mRNA sequence.

ACCESSION C18310  
 VERSION C18310  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 1 (bases 1 to 475)  
 AUTHORS Fujiwara,T., Hirano,H., Hishigaki,H., Horie,M., Kawai,A., Kuga,Y., Kyushiki,H., Nagata,M., Okuno,S., Ozaki,K., Shimizu,F., Shimada,Y., Shinomiya,H., Suzuki,M., Takaichi,A., Takeda,S., Watanabe,T., Maekawa,H., Nakamura,Y. and Takahashi,E.

TITLE Otsuka cDNA project  
 JOURNAL Unpublished (1996)  
 COMMENT On Sep 12, 1996 this sequence version replaced gi:1393837.  
 Contact: Tsutomu Fujiwara  
 Otsuka GEN Research Institute  
 Otsuka Pharmaceutical Co.,Ltd  
 463-10 Kagasuno Kawauchi-cho, Tokushima, 771-01 Japan  
 Tel: 0886-65-2888  
 Fax: 0886-37-1035.

FEATURES  
 Location/Qualifiers  
 1..475

/organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="GEN-560D07"  
 /clone\_lib="Human placenta cDNA (TFujiwara)"  
 /tissue\_type="placenta"

BASE COUNT 89 a 161 c 154 g 68 t 3 others  
 ORIGIN

alignment\_scores:

Quality: 51.00 Length: 10  
 Ratio: 5.100 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-08-653-294-24 x C18310

Align seg 1/1 to: C18310 from: 1 to: 475

1 ArgGluSerLeuArgAsnLeuArgGlyTyr 10  
 |||||  
 299 CGAGAGAGCTCGGAACCTCGCGGCTAC 328

seq\_name: gb\_est17:AA663896

seq\_documentation\_block:

LOCUS AA663896 479 bp mRNA EST 15-DEC-1997  
 DEFINITION ae74d01.s1 Stratagene schizo brain S11 Homo sapiens cDNA clone IMAGE:969889 3', similar to gb:M28203 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-62 B\*1504 ALPHA (HUMAN);, mRNA sequence.

ACCESSION AA663896  
 VERSION AA663896.1 GI:2617887  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 1 (bases 1 to 479)  
 AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,



Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M.,  
Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F.,  
Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.  
WashU-NCI human EST Project  
Unpublished (1997)

On Sep 12, 1996 this sequence version replaced gi:1394858.

Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@wustl.wustl.edu

This clone is available royalty-free through LLNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Possible reversed clone: polyt not found  
Seq primer: -40ml3 fwd. ET from Amersham  
High quality sequence stop: 430.

#### FEATURES

Source  
1. .479  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:96989"  
/clone\_lib="Stratagene schizo brain S11"  
/sex="male"  
/tissue\_type="schizophrenic brain S-11 frontal lobe"  
/dev\_stage="34 years old"  
/lab\_host="SOLR (kanamycin resistant)"  
/note="Vector: Bluescript SK-; Site 1: EcoRI; Library  
constructed from S-11 frontal lobe, male, 34 years old,  
50% caucasian, 50% Aleutian. Schizophrenic suicide.  
Random primed into EcoRI site of ZAP II Vector. Mass  
excised. Custom library. Avg insert length 1.4kb.  
Material obtained by Johnston N., Torrey, E.F., Yolken R.,  
and the Stanley Neuropathology Consortium - Analysis of  
RNAs from the Brains of Individuals with Psychiatric  
Diseases (Unpublished) Stanley Neurovirology Laboratory,  
Johns Hopkins School of Medicine, Baltimore MD."  
BASE COUNT 91 a 165 c 158 g 64 t 1 others  
ORIGIN

alignment\_scores:  
Quality: 51.00 Length: 10  
Ratio: 5.100 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

#### alignment\_block:

US-08-653-294-24 x AA663896 ..  
Align seg 1/1 to: AA663896 from: 1 to: 479

1 ArgGluserLeuArgAsnLeuArgGlyTyr 10  
|||||  
283 CGAGAGACCTGCGGACCTGCGGGCTAC 312

seq\_name: gb\_est35:AL039796

seq\_documentation\_block:  
LOCUS AL039796 503 bp mRNA EST 29-SEP-1999  
DEFINITION DKFp434B1912.1 434 (synonym: htes3) Homo sapiens cDNA clone  
DKFp434B1912 5', mRNA sequence.  
ACCESSION AL039796  
VERSION AL039796  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 503)

REFERENCE Duesterhoeft,A., Lauber,J., Mewes,H.W., Gassenhuber,J. and

Wiemann,S.

TITLE EST (Duesterhoeft, et al.)

JOURNAL Unpublished (1999)

#### COMMENT

On Jun 2, 1999 this sequence version replaced gi:4967270.  
Contact: Duesterhoeft A  
MPS  
Am Klopferpitz 18a D-82152 Martinsried, Germany  
This is the 5' sequence of the clone insert  
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
sequenced by Qiagen within the cDNA sequencing consortium of the  
German Genome Project.  
No sl sequence available.  
This clone is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

#### FEATURES

Source  
1. .503  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="DKFp434B1912"  
/clone\_lib="434 (synonym: htes3)"  
/tissue\_type="testis"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/note="Vector: pSport1; Site 1: NotI; Site 2: SalI"  
BASE COUNT 96 a 172 c 165 g 68 t 2 others  
ORIGIN

alignment\_scores:  
Quality: 51.00 Length: 10  
Ratio: 5.100 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

#### alignment\_block:

US-08-653-294-24 x AL039796 ..  
Align seg 1/1 to: AL039796 from: 1 to: 503  
1 ArgGluserLeuArgAsnLeuArgGlyTyr 10  
|||||  
300 CGAGAGACCTGCGGACCTGCGGGCTAC 329

seq\_name: gb\_est11:AA263135

seq\_documentation\_block:  
LOCUS AA263135 710 bp mRNA EST 02-JUL-1998  
DEFINITION PMY0598 KG1-a Lambda Zap Express cDNA library Homo sapiens cDNA 5',  
mRNA sequence.

ACCESSION AA263135  
VERSION AA263135.1 GI:1898941  
KEYWORDS EST.  
SOURCE human.

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 710)

AUTHORS Claudio,J.O., Liew,C.C., Dempsey,A.A., Cukerman,E., Stewart,A.K.,

Na,E., Atkins,H.I., Iscove,N.N. and Hawley,R.G.

Identification of sequence-tagged transcripts differentially

expressed within the human hematopoietic hierarchy

Genomics 50 (1), 44-52 (1998)

JOURNAL 98292493

MEDLINE

COMMENT On May 5, 1995 this sequence version replaced gi:797738.

Contact: Hawley RG

Oncology Research Laboratories

The Toronto Hospital

CRCS-424, 67 College St., Toronto, Ontario M5G 2M1, Canada

Tel: 416 3403834

Fax: 416 3403453

Email: r.hawley@utoronto.ca

Clone was randomly picked from KG1a primary library.

Seq primer: 5' GAATTAACCTCATTAAAGG 3'

High quality sequence stop: 710.

Location/Qualifiers

#### FEATURES



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 7, 2000, 18:04:37 ; Search time 111.22 Seconds  
(without alignments)  
4.241 Million cell updates/sec

Title: US-08-653-294-25  
Perfect score: 50  
Sequence: 1 RENLRTALRY 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 142080 seqs, 47169319 residues

Total number of hits satisfying chosen parameters: 142080

Minimum DB seq length: 0  
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database : PIR\_62: \*  
1: pir1: \*  
2: pir2: \*  
3: pir3: \*  
4: pir4: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	50	100.0	355	2 I37516	HLA-B alpha-chain - human (fragment)
2	50	100.0	362	2 S25415	class I histocompa
3	50	100.0	362	2 A45850	MHC class I histoc
4	50	100.0	362	2 I61861	MHC HLA-B*44.2 chai
5	50	100.0	362	2 I54442	MHC class I histoc
6	45	90.0	137	2 I80174	class I histocompa
7	45	90.0	359	1 HLBUB4	MHC class I histoc
8	45	90.0	362	2 I54457	MHC class I lympho
9	44	88.0	273	2 I38509	MHC class I histoc
10	44	88.0	274	2 I54463	MHC HLA-B*38 chain
11	44	88.0	354	2 I59308	class I histocompa
12	44	88.0	354	2 I80168	class I histocompa
13	44	88.0	354	2 I80167	class I histocompa
14	44	88.0	355	2 I80169	class I histocompa
15	44	88.0	355	2 I80171	class I histocompa
16	44	88.0	359	1 HLBUB2	MHC class I histoc
17	44	88.0	362	1 HLHUB8	MHC class I histoc
18	44	88.0	362	2 B30345	MHC class I histoc
19	44	88.0	362	2 JH0541	class I histocompa
20	44	88.0	362	2 JH0539	class I histocompa
21	44	88.0	362	2 JH0540	class I histocompa
22	44	88.0	362	2 A45834	MHC class I histoc
23	44	88.0	362	2 I84486	transmembrane glyco
24	44	88.0	362	2 I62045	gene HLA B-1517 pr
25	44	88.0	362	2 I84490	lymphocyte antigen
26	44	88.0	362	2 I37521	HLA-B*57.2 antigen
27	44	88.0	362	2 A30345	MHC class I histoc
28	44	88.0	362	2 I59633	MHC HLA-B histomem
29	44	88.0	362	2 S24434	class I histocompa
30	44	88.0	362	2 I37120	MHC class I histoc

31	44	88.0	363	2 S07113	class I histocompa
32	44	88.0	363	2 S03537	class I histocompa
33	44	88.0	364	2 D35997	MHC class I histoc
34	44	88.0	365	2 S77963	MHC class I histoc
35	44	88.0	365	2 I54416	HLA-AW24 protein -
36	44	88.0	365	2 I54493	MHC class I histoc
37	43	86.0	292	2 I57806	MHC H-2K-kml mRNA
38	43	86.0	362	2 B45876	class I histocompa
39	43	86.0	368	2 A60854	MHC class I histoc
40	43	86.0	368	2 I49712	H-2K-s - mouse
41	43	86.0	368	2 I49713	H-2K-sml - mouse
42	43	86.0	369	1 HLMSKK	MHC class I histoc
43	40	80.0	338	2 I56116	MHC HLA-B*27-HS - h
44	40	80.0	362	1 HLHUB2	MHC class I histoc
45	40	80.0	362	2 C35997	MHC class I histoc

## ALIGNMENTS

RESULT 1  
I37516  
HLA-B alpha-chain - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 04-Oct-1996 #sequence\_revision 04-Oct-1996 #text\_change 23-Jul-1999  
C:Accession: I37516  
R:Gauchat-Feliss, D.; Breur-Vriesendorp, B.S.; Rufer, N.; Jeannet, M.; Roosnek, E.; Ti  
Tissue Antigens 44, 261-264, 1994  
A:Title: Sequencing of a novel functional HLA-B\*44 subtype differing in two residues 1  
A:Reference number: I37516; MUID:95176328  
A:Accession: I37516  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-355 <RES>  
A:Cross-references: EMBL:X75953; NID:g791007; PIDN:CAA53566.1; PID:g791008  
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 100.0%; Score 50; DB 2; Length 355;  
Best Local Similarity 100.0%; Pred. No. 0.014;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RENLRTALRY 10  
|||||||  
Db 99 RENLRTALRY 108

RESULT 2  
S25415  
class I histocompatibility antigen HLA-B\*4403 alpha chain - human  
C:Species: Homo sapiens (man)  
C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 23-Jul-1999  
C:Accession: S25415  
R:Fleischhauser, K.; Kernan, N.A.; Dupont, B.; Yang, S.Y.  
Tissue Antigens 37, 133-137, 1991  
A:Title: The two major subtypes of HLA-B\*44 differ for a single amino acid in codon 15  
A:Reference number: S25415; MUID:91335451  
A:Accession: S25415  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-362 <FLE>  
A:Cross-references: EMBL:X64366; NID:g32178; PIDN:CAA45718.1; PID:g32179  
C:Genetics:  
A:Gene: GBB:HLA-B  
A:Cross-references: GDB:120048; OMIM:142830  
A:Map position: 6p21.3-6p21.3  
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology  
C:Keywords: transmembrane protein  
F:220-285/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 50; DB 2; Length 362;  
Best Local Similarity 100.0%; Pred. No. 0.014;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RENLRTALRY 10

Db 99 RENLRTALRY 108

RESULT 3

A45850

MHC class I histocompatibility antigen HLA-B13.1 - human

C:Species: Homo sapiens (man)

C>Date: 03-Jun-1993 #sequence\_revision 03-Jun-1993 #text\_change 23-Jul-1999

C:Accession: A45850

R:Kato, K.; Dupont, B.; Yang, S.Y.

A:Title: Localization of nucleotide sequence which determines mongoloid subtype of HLA-B

A:Reference number: A45850; MUID:89122134

A:Accession: A45850

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-362 <RAT>

A:Cross-references: GB:M24075; NID:g187703; PIDN:AAA59627.1; PID:g386884

C:Genetics:

A:Gene: GDB:HLA-B

A:Cross-references: GDB:120048; OMIM:142830

A:Map position: 6p21.3-6p21.3

C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

F:220-285/Domain: immunoglobulin homology <IMM>

Query Match

Best Local Similarity 100.0%; Score 50; DB 2; Length 362;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RENLRTALRY 10

Db 99 RENLRTALRY 108

RESULT 4

I61861

MHC HLA-B44.2 chain - human

C:Species: Homo sapiens (man)

C>Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 23-Jul-1999

C:Accession: I61861

R:Parham, P.; Lawlor, D.A.; Lomen, C.E.; Ennis, P.D.

J. Immunol. 142, 3937-3950, 1989

A:Title: Diversity and diversification of HLA-A,B,C alleles.

A:Reference number: I36956; MUID:89235215

A:Accession: I61861

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-362 <RES>

A:Cross-references: GB:M24038; NID:g187811; PIDN:AAA59663.1; PID:g386900

C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match

Best Local Similarity 100.0%; Score 50; DB 2; Length 362;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RENLRTALRY 10

Db 99 RENLRTALRY 108

RESULT 5

I54442

MHC class I histocompatibility antigen HLA-B13 precursor - human

C:Species: Homo sapiens (man)

C>Date: 07-Jun-1996 #sequence\_revision 07-Jun-1996 #text\_change 23-Jul-1999

C:Accession: I54442; I61858

R:Zemmour, J.; Ennis, P.D.; Parham, P.; Dupont, B.

Immunogenetics 27, 281-287, 1988  
A:Title: Comparison of the structure of HLA-B\*47 to HLA-B13 and its relationship to 2  
A:Reference number: I54442; MUID:88152906

A:Accession: I54442

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-362 <ZEM>

A:Cross-references: GB:M19757; NID:g184161; PIDN:AAA52657.1; PID:g386774

R:Parham, P.; Lawlor, D.A.; Lomen, C.E.; Ennis, P.D.

J. Immunol. 142, 3937-3950, 1989

A:Title: Diversity and diversification of HLA-A,B,C alleles.

A:Reference number: I36956; MUID:89235215

A:Accession: I61858

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-362 <PAR>

A:Cross-references: GB:M24041; NID:g187805; PIDN:AAA59660.1; PID:g386897

C:Genetics:

A:Gene: GDB:HLA-B

A:Cross-references: GDB:120048; OMIM:142830

A:Map position: 6p21.3-6p21.3

C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match

Best Local Similarity 100.0%; Score 50; DB 2; Length 362;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RENLRTALRY 10

Db 99 RENLRTALRY 108

RESULT 6

I80174

class I histocompatibility antigen - chimpanzee (fragment)

C:Species: Pan troglodytes (chimpanzee)

C>Date: 24-May-1996 #sequence\_revision 24-May-1996 #text\_change 23-Jul-1999

C:Accession: I80174

R:McAdam, S.N.; Boyson, J.E.; Liu, X.; Garber, T.L.; Hughes, A.L.; Bontrop, R.E.; Wat

Proc. Natl. Acad. Sci. U.S.A. 91, 5893-5897, 1994

A:Title: A uniquely high level of recombination at the HLA-B locus.

A:Reference number: I59308; MUID:94286344

A:Accession: I80174

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-137 <RES>

A:Cross-references: EMBL:U05585; NID:g454787; PIDN:AAA50188.1; PID:g454788

C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match

Best Local Similarity 90.0%; Score 45; DB 2; Length 137;

Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RENLRTALRY 10

Db 40 RENLRTALRY 49

RESULT 7

HLHUB4

MHC class I histocompatibility antigen HLA-B44 alpha chain precursor - human (fragmen

C:Species: Homo sapiens (man)

C>Date: 31-Dec-1988 #sequence\_revision 31-Dec-1988 #text\_change 22-Jun-1999

C:Accession: A25295

R:Kottmann, A.H.; Seemann, G.H.A.; Guesow, H.D.; Roos, M.H.

Immunogenetics 23, 396-400, 1986

A:Title: DNA sequence of the coding region of the HLA-B44 gene.

A:Reference number: A25295; MUID:86249389

A:Accession: A25295

A:Molecule type: mRNA

A:Residues: 1-359 <KOT>

A:Cross-references: GB:M15470; NID:g187680; PIDN:AAA59619.1; PID:g386883  
C:Genetics:

A:Gene: GDB:HLA-B  
A:Cross-references: GDB:120048; OMIM:142830  
A:Map position: 6p21.3-6p21.3  
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology  
C:Keywords: duplication; glycoprotein; heterodimer; transmembrane protein; transplantati  
F:1-21/Domain: signal sequence (fragment) #status predicted <SIG>  
F:22-359/Product: class I histocompatibility antigen HLA-B44 alpha chain #status predict  
F:22-304/Domain: extracellular #status predicted <EXT>  
F:22-111/Domain: alpha-1 <EX1>  
F:112-203/Domain: alpha-2 <EX2>  
F:217-282/Domain: immunoglobulin homology <IMM>  
F:305-328/Domain: transmembrane #status predicted <TM>  
F:329-359/Domain: intracellular #status predicted <INT>  
F:107/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 90.0%; Score 45; DB 1; Length 359;  
Best Local Similarity 90.0%; Pred. No. 0.14;  
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RENVLTALRY 10  
||||| ||  
DB 96 RENVLTAARY 105

RESULT 8  
I54457

MHC class I lymphocyte antigen - human  
C:Species: Homo sapiens (man)  
C:Date: 07-Jun-1996 #sequence\_revision 07-Jun-1996 #text\_change 23-Jul-1999  
C:Accession: I54457  
R:Pohla, H.; Kuon, W.; Tabaczewski, P.; Doerner, C.; Weiss, E.H.  
Immunogenetics 29, 297-307, 1989  
A:Title: Allelic variation in HLA-B and HLA-C sequences and the evolution of the HLA-B a  
A:Reference number: I54457; MUID:89233295  
A:Accession: I54457  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-362 <RES>

A:Cross-references: GB:M28205; NID:g576470; PIDN:AAA57145.1; PID:g576471  
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 90.0%; Score 45; DB 2; Length 362;  
Best Local Similarity 90.0%; Pred. No. 0.14;  
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RENVLTALRY 10  
||||| ||  
DB 99 RENVLTALRY 108

RESULT 9  
I38509

MHC class I histocompatibility antigen - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 06-Sep-1996 #sequence\_revision 06-Sep-1996 #text\_change 23-Jul-1999  
C:Accession: I38509  
R:Cereb, N.; Choi, J.W.; Riu, K.Z.; Yang, S.Y.  
Tissue Antigens 44, 271-273, 1994  
A:Title: HLA-B\*5105, a newly identified B51 IEF variant.  
A:Reference number: I38509; MUID:95176331  
A:Accession: I38509

A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-273 <RES>

A:Cross-references: EMBL:U06697; NID:g469544; PIDN:AAA92997.1; PID:g469545  
C:Genetics:

A:Gene: GDB:HLA-B  
A:Cross-references: GDB:120048; OMIM:142830  
A:Map position: 6p21.3-6p21.3

C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 88.0%; Score 44; DB 2; Length 273;  
Best Local Similarity 90.0%; Pred. No. 0.17;  
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RENVLTALRY 10  
||||| ||  
DB 74 RENVLTALRY 83

RESULT 10  
I54463

MHC HLA-B38 chain - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 07-Jun-1996 #sequence\_revision 07-Jun-1996 #text\_change 23-Jul-1999  
C:Accession: I54463  
R:Mueller, C.A.; Engler-Blum, G.; Gekeler, V.; Steiert, I.; Weiss, E.; Schmidt, H.  
Immunogenetics 30, 200-207, 1989  
A:Title: Genetic and serological heterogeneity of the supertypic HLA-B locus specific  
A:Reference number: I54463; MUID:89379286  
A:Accession: I54463  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-274 <RES>  
A:Cross-references: GB:M29864; NID:g187674; PIDN:AAA36222.1; PID:g187675  
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 88.0%; Score 44; DB 2; Length 274;  
Best Local Similarity 90.0%; Pred. No. 0.17;  
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RENVLTALRY 10  
||||| ||  
DB 75 RENVLTALRY 84

RESULT 11  
I59308

class I histocompatibility antigen - pygmy chimpanzee (fragment)  
C:Species: Pan paniscus (pygmy chimpanzee, bonobo)  
C:Date: 31-May-1996 #sequence\_revision 31-May-1996 #text\_change 23-Jul-1999  
C:Accession: I59308  
R:McAdam, S.N.; Boyson, J.E.; Liu, X.; Garber, T.L.; Hughes, A.L.; Bontrop, R.E.; Wat  
Proc. Natl. Acad. Sci. U.S.A. 91, 5893-5897, 1994  
A:Title: A uniquely high level of recombination at the HLA-B locus.  
A:Reference number: I59308; MUID:94286544  
A:Accession: I59308

A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-354 <RES>

A:Cross-references: EMBL:U05575; NID:g454767; PIDN:AAA50178.1; PID:g454768  
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 88.0%; Score 44; DB 2; Length 354;  
Best Local Similarity 90.0%; Pred. No. 0.22;  
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RENVLTALRY 10  
||||| ||  
DB 91 RENVLTALRY 100

RESULT 12  
I80168

class I histocompatibility antigen - chimpanzee (fragment)  
C:Species: Pan troglodytes (chimpanzee)

C:Date: 24-May-1996 #sequence\_revision 24-May-1996 #text\_change 23-Jul-1999  
C:Accession: I80168  
R:McAdam, S.N.; Boyson, J.E.; Liu, X.; Garber, T.L.; Hughes, A.L.; Bontrop, R.E.; Wat

Proc. Natl. Acad. Sci. U.S.A. 91, 5893-5897, 1994  
A:Title: A uniquely high level of recombination at the HLA-B locus.  
A:Reference number: I59308; MUID:94286544

A:Accession: I80168  
A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: mRNA

A:Residues: 1-354 <RES>

A:Cross-references: EMBL:U05579; NID:q454775; PIDN:AAA50182.1; PID:q454776  
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 88.0%; Score 44; DB 2; Length 354;  
Best Local Similarity 90.0%; Pred. No. 0.22;  
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RENLRLALRY 10  
||||| |||||  
Db 91 RENLRLALRY 100

## RESULT 13

I80167  
class I histocompatibility antigen - pygmy chimpanzee (fragment)

C:Species: Pan paniscus (pygmy chimpanzee, bonobo)

C:Date: 24-May-1996 #sequence\_revision 24-May-1996 #text\_change 23-Jul-1999

C:Accession: I80167

R:McAdam, S.N.; Boyson, J.E.; Liu, X.; Garber, T.L.; Hughes, A.L.; Bontrop, R.E.; Watkin

Proc. Natl. Acad. Sci. U.S.A. 91, 5893-5897, 1994

A:Title: A uniquely high level of recombination at the HLA-B locus.

A:Reference number: I59308; MUID:94286544

A:Accession: I80167

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: mRNA

A:Residues: 1-354 <RES>

A:Cross-references: EMBL:U05578; NID:q454773; PIDN:AAA50181.1; PID:q454774

C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 88.0%; Score 44; DB 2; Length 354;  
Best Local Similarity 90.0%; Pred. No. 0.22;  
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RENLRLALRY 10  
||||| |||||  
Db 91 RENLRLALRY 100

## RESULT 14

I80169

class I histocompatibility antigen - chimpanzee (fragment)

C:Species: Pan troglodytes (chimpanzee)

C:Date: 24-May-1996 #sequence\_revision 24-May-1996 #text\_change 23-Jul-1999

C:Accession: I80169

R:McAdam, S.N.; Boyson, J.E.; Liu, X.; Garber, T.L.; Hughes, A.L.; Bontrop, R.E.; Watkin

Proc. Natl. Acad. Sci. U.S.A. 91, 5893-5897, 1994

A:Title: A uniquely high level of recombination at the HLA-B locus.

A:Reference number: I59308; MUID:94286544

A:Accession: I80169

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: mRNA

A:Residues: 1-355 <RES>

A:Cross-references: EMBL:U05580; NID:q454777; PIDN:AAA50183.1; PID:q454778

C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 88.0%; Score 44; DB 2; Length 355;  
Best Local Similarity 90.0%; Pred. No. 0.22;  
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RENLRLALRY 10  
||||| |||||  
Db 91 RENLRLALRY 100

## RESULT 15

I80171

class I histocompatibility antigen - chimpanzee (fragment)

C:Species: Pan troglodytes (chimpanzee)

C:Date: 24-May-1996 #sequence\_revision 24-May-1996 #text\_change 23-Jul-1999

C:Accession: I80171

R:McAdam, S.N.; Boyson, J.E.; Liu, X.; Garber, T.L.; Hughes, A.L.; Bontrop, R.E.; Wat

Proc. Natl. Acad. Sci. U.S.A. 91, 5893-5897, 1994

A:Title: A uniquely high level of recombination at the HLA-B locus.

A:Reference number: I59308; MUID:94286544

A:Accession: I80171

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: mRNA

A:Residues: 1-355 <RES>

A:Cross-references: EMBL:U05582; NID:q454781; PIDN:AAA50185.1; PID:q454782

C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 88.0%; Score 44; DB 2; Length 355;  
Best Local Similarity 90.0%; Pred. No. 0.22;  
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RENLRLALRY 10  
||||| |||||  
Db 91 RENLRLALRY 100

Search completed: February 7, 2000, 18:04:37

Job time: 22203 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2000, 01:25:58 ; Search time 68.63 seconds  
(without alignments)  
4.352 Million cell updates/sec

Title: US-08-653-294-25  
Perfect score: 50  
Sequence: 1 RENTRYALRY 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 82229 seqs, 29864866 residues

Total number of hits satisfying chosen parameters: 82229

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database : SwissProt\_38:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Query No.	Score	Match	Length	DB ID	Description
1	50	100.0	362	1	1B05_HUMAN	P30461 homo sapien
2	50	100.0	362	1	1B41_HUMAN	P30481 homo sapien
3	50	100.0	362	1	1B42_HUMAN	P30482 homo sapien
4	45	90.0	359	1	1B40_HUMAN	P10320 homo sapien
5	44	88.0	359	1	1B01_PANTR	P13750 pan troglod
6	44	88.0	362	1	1B01_GORGO	P30379 gorilla gor
7	44	88.0	362	1	1B02_GORGO	P30380 gorilla gor
8	44	88.0	362	1	1B03_GORGO	P30381 gorilla gor
9	44	88.0	362	1	1B15_HUMAN	P10317 homo sapien
10	44	88.0	362	1	1B47_HUMAN	P30487 homo sapien
11	44	88.0	362	1	1B49_HUMAN	P18464 homo sapien
12	44	88.0	362	1	1B52_HUMAN	P30489 homo sapien
13	44	88.0	362	1	1B53_HUMAN	P30490 homo sapien
14	44	88.0	362	1	1B54_HUMAN	P30491 homo sapien
15	44	88.0	362	1	1B60_HUMAN	P18465 homo sapien
16	44	88.0	362	1	1B61_HUMAN	P30497 homo sapien
17	44	88.0	362	1	1B62_HUMAN	P10319 homo sapien
18	44	88.0	362	1	HLAH_HUMAN	P01893 homo sapien
19	44	88.0	365	1	1A23_HUMAN	P30447 homo sapien
20	44	88.0	365	1	1A24_HUMAN	P05534 homo sapien
21	43	86.0	369	1	HA1K_MOUSE	P04223 mus musculu
22	40	80.0	338	1	1B20_HUMAN	P30467 homo sapien
23	40	80.0	361	1	1B14_HUMAN	P03989 homo sapien
24	40	80.0	362	1	1B16_HUMAN	P19373 homo sapien
25	40	80.0	362	1	1B18_HUMAN	P10318 homo sapien
26	40	80.0	362	1	1B19_HUMAN	Q08136 homo sapien
27	40	80.0	362	1	1B29_HUMAN	P18463 homo sapien
28	40	80.0	362	1	1B45_HUMAN	P30485 homo sapien
29	39	78.0	365	1	1A25_HUMAN	P18462 homo sapien
30	39	78.0	365	1	1A32_HUMAN	P10314 homo sapien
31	38	76.0	298	1	HA1Y_MOUSE	P01895 mus musculu
32	38	76.0	361	1	HA1A_RABIT	P01894 cryptolagus
33	38	76.0	361	1	HA1B_RABIT	P06140 cryptolagus
34	38	76.0	365	1	HA12_MOUSE	P01900 mus musculu

## RESULT 1

1B05\_HUMAN  
ID 1B05\_HUMAN STANDARD; PRT; 362 AA.  
AC P30461;  
DT 01-APR-1993 (Rel. 25, Created)  
DT 01-APR-1993 (Rel. 25, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-13 B\*1301 ALPHA CHAIN  
DE PRECURSOR (B13.1).  
GN HLA-B OR HLAB.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA MEDLINE; 89235215.  
RX PARHAM P., LAWOR D.A., LOMEN C.E., ENNIS P.D.;  
RT "Diversity and diversification of HLA-A,B,C alleles.";  
RL J. Immunol. 142:3937-3950(1989).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA MEDLINE; 88152906.  
RX ZEMMOUR J., ENNIS P.D., PARHAM P., DUPONT B.;  
RT "Comparison of the structure of HLA-B\*47 to HLA-B13 and its  
RT relationship to 21-hydroxylase deficiency.";  
RN Immunogenetics 27:281-287(1988).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE-BLOOD;  
RX MEDLINE; 96053518.  
RA LIN L., TOKUNAGA K., NAKAJIMA F., ISHIKAWA Y., KASHIWASE K.,  
RA TANAKA H., KUNATA S., SIDELTSEVA E., AKAZA T., TADOKORO K.,  
RA SHIBATA Y., CHANDANAYINGYONG D., JUJI T.;  
RT "Both HLA-B\*1301 and B\*1302 exist in Asian populations and are  
RT associated with different haplotypes.";  
RL Hum. Immunol. 43:51-56(1995).  
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO  
CC THE IMMUNE SYSTEM.  
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
CC MICROGLOBULIN).  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; M24041; AAA59660.1; -;  
CC EMBL; M19757; AAA52657.1; -;  
CC EMBL; D50291; BAA08822.1; -;  
CC HSSP; P30491; lalm.  
CC MIM; 142830; -;  
CC PROSITE; PS00290; IG\_MHC; 1.  
CC PFAM; PF00047; Ig; 1.  
CC PFAM; PF00129; MHC\_I; 1.

35 35 70.0 555 1 POST\_XENLA P27536 xenopus lae  
36 34 68.0 361 1 HA1U\_MOUSE P14433 mus musculu  
37 34 68.0 365 1 1A04\_GORGO P30378 gorilla gor  
38 34 68.0 384 1 HA1T\_MOUSE P14432 mus musculu  
39 34 68.0 582 1 MNT\_HUMAN Q9583 homo sapien  
40 34 68.0 591 1 MNT\_MOUSE Q08789 mus musculu  
41 33 66.0 225 1 SODM\_BACSU P54375 bacillus su  
42 33 66.0 267 1 YADB\_SCHPO Q09836 schizosacch  
43 33 66.0 328 1 HA1Q\_MOUSE P14428 mus musculu  
44 33 66.0 362 1 HA13\_MOUSE P14426 mus musculu  
45 33 66.0 368 1 HA1W\_MOUSE P03991 mus musculu

## ALIGNMENTS

KW MHC I; Transmembrane; Glycoprotein; Signal.  
 FT SIGNAL 1 24  
 FT CHAIN 25 362 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,  
 FT B-13 B\*1301 ALPHA CHAIN.  
 FT DOMAIN 25 114  
 FT DOMAIN 115 206 EXTRACELLULAR ALPHA-1.  
 FT DOMAIN 207 298 EXTRACELLULAR ALPHA-2.  
 FT DOMAIN 299 309 EXTRACELLULAR ALPHA-3.  
 FT TRANSMEM 310 333 CONNECTING PEPTIDE.  
 FT DOMAIN 334 362 CYTOPLASMIC TAIL.  
 FT CARBOHYD 110 110 BY SIMILARITY.  
 FT DISULFID 125 188 BY SIMILARITY.  
 FT DISULFID 227 283 BY SIMILARITY.  
 SQ SEQUENCE 362 AA; 40474 MW; 28B67875 CRC32;

Query Match 100.0%; Score 50; DB 1; Length 362;  
 Best Local Similarity 100.0%; Pred. No. 0.0042;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RENLRTALRY 10  
 (|||||)  
 Db 99 RENLRTALRY 108

RESULT 2  
 1B41\_HUMAN  
 ID 1B41\_HUMAN STANDARD; PRT; 362 AA.  
 AC P30481;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, BW-44(B-12) B\*4402 ALPHA CHAIN  
 DE PRECURSOR (B44.2).  
 GN HLA-B OR HLAB.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 89235215.  
 RA PARHAM P.; LAWOR D.A.; LOMEN C.E.; ENNIS P.D.;  
 RT "Diversity and diversification of HLA-A,B,C alleles."  
 RL J. Immunol. 142:3937-3950(1989).  
 CC -1- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO  
 THE IMMUNE SYSTEM.  
 CC -1- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
 MICROGLOBULIN).  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL; M24038; AAB59663.1;  
 CC HSP; P30491; IALM.  
 CC MIM; 142830;  
 CC PROSITE; PS00290; IG\_MHC; 1.  
 CC PFAM; PF00047; ig; 1.  
 CC PFAM; PF00129; MHC.I; 1.  
 KW MHC I; Transmembrane; Glycoprotein; Signal.  
 FT SIGNAL 1 24  
 FT CHAIN 25 362 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,  
 FT BW-44(B-12) B\*4402 ALPHA CHAIN.  
 FT DOMAIN 25 114  
 FT DOMAIN 115 206 EXTRACELLULAR ALPHA-1.  
 FT DOMAIN 207 298 EXTRACELLULAR ALPHA-2.  
 FT DOMAIN 299 309 EXTRACELLULAR ALPHA-3.  
 FT TRANSMEM 310 333 CONNECTING PEPTIDE.  
 FT DOMAIN 334 362 CYTOPLASMIC TAIL.

FT CARBOHYD 110 110 BY SIMILARITY.  
 FT DISULFID 125 188 BY SIMILARITY.  
 FT DISULFID 227 283 BY SIMILARITY.  
 SQ SEQUENCE 362 AA; 40481 MW; D0AE6DD5 CRC32;

Query Match 100.0%; Score 50; DB 1; Length 362;  
 Best Local Similarity 100.0%; Pred. No. 0.0042;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RENLRTALRY 10  
 (|||||)  
 Db 99 RENLRTALRY 108

RESULT 3  
 1B42\_HUMAN  
 ID 1B42\_HUMAN STANDARD; PRT; 362 AA.  
 AC P30482;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, BW-44(B-12) B\*4403 ALPHA CHAIN  
 DE PRECURSOR.  
 GN HLA-B OR HLAB.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 91335451.  
 RA FLEISCHCHAUER K.; KERNAN N.A.; DUPONT B.; YANG S.Y.;  
 RT "The two major subtypes of HLA-B\*44 differ for a single amino acid in  
 RT codon 156."  
 RL Tissue Antigens 37:133-137(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 96435470.  
 RA ADAMS E.J.; LITTLE A.M.; ARNETT K.L.; MCAULEY J.E.; WILLIAMS R.C.;  
 RA PARHAM P.;  
 RT "Three new HLA-B alleles found in Mexican-Americans."  
 RL Tissue Antigens 46:414-416(1995).  
 CC -1- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO  
 THE IMMUNE SYSTEM.  
 CC -1- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
 MICROGLOBULIN).  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL; X64366; CAA45718.1;  
 CC EMBL; L42282; AAB51454.1;  
 CC EMBL; L42283; AAB51455.1;  
 CC PIR; S25415; S25415.  
 CC HSP; P30491; IALM.  
 CC MIM; 142830;  
 CC PROSITE; PS00290; IG\_MHC; 1.  
 CC PFAM; PF00047; ig; 1.  
 CC PFAM; PF00129; MHC.I; 1.  
 KW MHC I; Transmembrane; Glycoprotein; Signal.  
 FT SIGNAL 1 24  
 FT CHAIN 25 362 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,  
 FT BW-44(B-12) B\*4403 ALPHA CHAIN.  
 FT DOMAIN 25 114  
 FT DOMAIN 115 206 EXTRACELLULAR ALPHA-1.  
 FT DOMAIN 207 298 EXTRACELLULAR ALPHA-2.  
 FT DOMAIN 299 309 EXTRACELLULAR ALPHA-3.  
 FT TRANSMEM 310 333 CONNECTING PEPTIDE.



FT DOMAIN 334 362 CYTOPLASMIC TAIL.  
FT CARBOHYD 110 110 BY SIMILARITY.  
FT DISULFID 125 188 BY SIMILARITY.  
FT DISULFID 227 283 BY SIMILARITY.  
SQ SEQUENCE 362 AA; 40479 MW; E72CE869 CRC32;

Query Match 100.0%; Score 50; DB 1; Length 362;  
Best Local Similarity 100.0%; Pred. No. 0.0042;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 RENLTRLRY 10  
Db 99 RENLTRLRY 108

RESULT 4  
ID 1B40\_HUMAN STANDARD; PRT: 359 AA.  
AC P10320;  
DT 01-MAR-1989 (Rel. 10, Created)  
DT 01-MAR-1989 (Rel. 10, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, BW-44(B-12) B\*4401 ALPHA CHAIN  
DE PRECURSOR (B44.1) (FRAGMENT).  
GN HLA-B OR HLAB.

OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE; 86249389.

RA KOTTMANN A.H., SEEMANN G.H.A., GUESSOW H.D., ROOS M.H.;

RT "DNA sequence of the coding region of the HLA-B44 gene.";

RL Immunogenetics 23:396-400(1986).

CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO

CC THE IMMUNE SYSTEM.

CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-

CC MICROGLOBULIN).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; M15470; AAA59619.1; -.

DR PIR; A25295; HLHUB4.

DR HSSP; P30491; 1ALM.

DR MIM; 142830; -.

DR PROSITE; PS00290; IG\_MHC; 1.

DR PFAM; PF00047; Ig; 1.

DR PFAM; PF00129; MHC\_I; 1.

KW MHC I; Transmembrane; Glycoprotein; Signal.

FT NON\_TER 1 1

FT SIGNAL <1 21

FT CHAIN 22 359

FT HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,

FT BW-44 (B-12) B\*4401 ALPHA CHAIN.

FT EXTRACELLULAR ALPHA-1.

FT EXTRACELLULAR ALPHA-2.

FT EXTRACELLULAR ALPHA-3.

FT CONNECTING PEPTIDE.

FT CYTOPLASMIC.

FT BY SIMILARITY.

FT BY SIMILARITY.

FT BY SIMILARITY.

FT BY SIMILARITY.

FT BY SIMILARITY.

FT BY SIMILARITY.

FT BY SIMILARITY.

FT BY SIMILARITY.

FT BY SIMILARITY.

FT BY SIMILARITY.

FT BY SIMILARITY.

FT BY SIMILARITY.

FT BY SIMILARITY.

Best Local Similarity 90.0%; Pred. No. 0.047;  
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Oy 1 RENLTRLRY 10  
Db 96 RENLTRLRY 105

RESULT 5

ID 1B01\_PANTR

AC P13750;

DT 01-JAN-1990 (Rel. 13, Created)

DT 01-JAN-1990 (Rel. 13, Last sequence update)

DT 01-APR-1993 (Rel. 25, Last annotation update)

DE CHLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-1 ALPHA CHAIN PRECURSOR

DE (FRAGMENT).

OS Pan troglodytes (Chimpanzee).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

OC Eutheria; Primates; Catarrhini; Hominidae; Pan.

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE; 89030641.

RA MAYER W.E., JONKER M., KLEIN D., IVANYI P., VAN SEVENTER G.,

RA KLEIN J.;

RT "Nucleotide sequences of chimpanzee MHC class I alleles: evidence for

RT trans-species mode of evolution.";

RL EMBO J. 7:2765-2774(1988).

RN [2]

RP REVISIONS.

RA MAYER W.;

RL Submitted (FEB-1989) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO

CC THE IMMUNE SYSTEM.

CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-

CC MICROGLOBULIN).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; X13115; GAA31507.1; -.

DR PIR; S03537; S03537.

DR HSSP; P03989; 1HSA.

DR PROSITE; PS00290; IG\_MHC; 1.

DR PFAM; PF00047; Ig; 1.

DR PFAM; PF00129; MHC\_I; 1.

KW MHC I; Transmembrane; Glycoprotein; Signal.

FT NON\_TER 1 1

FT SIGNAL <1 20

FT CHAIN 21 359

FT CHLA CLASS I HISTOCOMPATIBILITY ANTIGEN,

FT B-1 ALPHA CHAIN.

FT EXTRACELLULAR ALPHA-1.

FT EXTRACELLULAR ALPHA-2.

FT EXTRACELLULAR ALPHA-3.

FT CONNECTING PEPTIDE.

FT CYTOPLASMIC TAIL.

FT BY SIMILARITY.

FT BY SIMILARITY.

FT BY SIMILARITY.

FT BY SIMILARITY.

FT BY SIMILARITY.

FT BY SIMILARITY.

FT BY SIMILARITY.

FT BY SIMILARITY.

FT BY SIMILARITY.

FT BY SIMILARITY.

FT BY SIMILARITY.

FT BY SIMILARITY.

FT BY SIMILARITY.

FT BY SIMILARITY.

FT BY SIMILARITY.

FT BY SIMILARITY.

FT BY SIMILARITY.

FT BY SIMILARITY.

FT BY SIMILARITY.

FT BY SIMILARITY.

FT BY SIMILARITY.

FT BY SIMILARITY.

Query Match 88.0%; Score 44; DB 1; Length 359;  
Best Local Similarity 90.0%; Pred. No. 0.075;  
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Oy 1 RENLTRLRY 10

```

Query Match      88.0%; Score 44; DB 1; Length 362;
Best Local Similarity 90.0%; Pred. No. 0.076;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY      1 RENVLTALRY 10
        ||||| ||||
Db.      99 RENVLTALRY 108

RESULT      7
1B02_GORGO
ID      1B02_GORGO      STANDARD;      PRT;      362 AA.
AC      P30380;
DT      01-APR-1993 (Rel. 25, Created)
DT      01-APR-1993 (Rel. 25, Last sequence update)
DT      01-APR-1993 (Rel. 25, Last annotation update)

```

CC -1- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO  
 CC THE IMMUNE SYSTEM.  
 CC -1- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
 CC MICROGLOBULIN).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; X60254; CAA42806.1; -  
 DR PIR; JH0541; JH0541.  
 DR HSP; P03989; ILSA.  
 DR PROSITE; PS00290; IG\_MHC; 1.  
 DR PFAM; PF00047; Ig; 1.  
 DR MHC I; Transmembrane; Glycoprotein; Signal.  
 FT SIGNAL 1 24  
 FT CHAIN 25 362  
 FT CLASS I HISTOCOMPATIBILITY ANTIGEN,  
 FT GOGO-B0103 ALPHA CHAIN.  
 FT DOMAIN 25 114  
 FT DOMAIN 115 206  
 FT DOMAIN 207 298  
 FT DOMAIN 299 308  
 FT DOMAIN 309 332  
 FT TRANSMEM 309 332  
 FT DOMAIN 333 362  
 FT DISULFID 125 188  
 FT DISULFID 227 283  
 FT CARBOHYD 110 110  
 FT SEQUENCE 362 AA; 40248 MW; FEAGA941 CRC32;  
 SQ

Query Match 88.0%; Score 44; DB 1; Length 362;  
 Best Local Similarity 90.0%; Pred. No. 0.076;  
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RENLRYALRY 10  
 ||||| |||||  
 Db 99 RENLRYALRY 108

RESULT 9  
 ID 1B15\_HUMAN STANDARD; PRT; 362 AA.  
 AC P10317;  
 DT 01-MAR-1989 (Rel. 10, Created)  
 DT 01-MAR-1989 (Rel. 10, Last sequence update)  
 DT 01-APR-1993 (Rel. 25, Last annotation update)  
 DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-27 B\*2702 ALPHA CHAIN  
 DE PRECURSOR (B-27K) (B27.2).  
 GN HLA-B OR HLAB.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 86220133.  
 RA SEEMANN G.H.A., REIN R.S., BROWN C.S., PLOEGH H.L.;  
 RT "Gene conversion-like mechanisms may generate polymorphism in human  
 RT class I genes".  
 RL EMBO J. 5:547-552(1986).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA PARHAM P., ARNETT K.L., ADAMS E.J.;  
 RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE OF 86-107 AND 171-181.  
 RX MEDLINE; 86042671.  
 RA VEGA M.A., EZQUEIRA A., ROJO S., APARICIO P., BRAGADO R.,  
 RA LOPEZ DE CASTRO J.A.;

RT \*Structural analysis of an HLA-B27 functional variant: identification  
 RT of residues that contribute to the specificity of recognition by  
 RT cytolytic T lymphocytes\*;  
 RL Proc. Natl. Acad. Sci. U.S.A. 82:7394-7398(1985).  
 CC -1- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO  
 CC THE IMMUNE SYSTEM.  
 CC -1- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
 CC MICROGLOBULIN).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; X03664; CAA27301.1; -  
 DR EMBL; X03667; CAA27301.1; JOINED.  
 DR EMBL; L38504; AAA69724.1; -  
 DR PIR; B25092; HLHURK.  
 DR HSP; P03989; ILSA.  
 DR MIM; 142830; -  
 DR PROSITE; PS00290; IG\_MHC; 1.  
 DR PFAM; PF00047; Ig; 1.  
 DR PFAM; PF00129; MHC I; 1.  
 KW MHC I; Transmembrane; Glycoprotein; Signal.  
 FT SIGNAL 1 24  
 FT CHAIN 25 362  
 FT HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,  
 FT B-27 B\*2702 ALPHA CHAIN.  
 FT DOMAIN 25 114  
 FT DOMAIN 115 206  
 FT DOMAIN 207 298  
 FT DOMAIN 299 308  
 FT TRANSMEM 309 332  
 FT DOMAIN 333 362  
 FT CARBOHYD 110 110  
 FT DISULFID 125 188  
 FT DISULFID 227 283  
 FT SEQUENCE 362 AA; 40397 MW; 9798F0BB CRC32;  
 SQ

Query Match 88.0%; Score 44; DB 1; Length 362;  
 Best Local Similarity 90.0%; Pred. No. 0.076;  
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RENLRYALRY 10  
 ||||| |||||  
 Db 99 RENLRYALRY 108

RESULT 10  
 ID 1B47\_HUMAN STANDARD; PRT; 362 AA.  
 AC P30487;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 01-FEB-1996 (Rel. 33, Last annotation update)  
 DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-49(B-21) B\*4901 ALPHA CHAIN  
 DE PRECURSOR.  
 GN HLA-B OR HLAB.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 89235215.  
 RA PARHAM P., LAWOR D.A., LOMEN C.E., ENNIS P.D.;  
 RT "Diversity and diversification of HLA-A,B,C alleles".  
 RL J. Immunol. 142:3937-3950(1989).  
 RN [2]  
 RP REVISION TO 78.  
 RX MEDLINE; 93056529.

```

RA HILDEBRAND W.H., MADRIGAL J.A., BELICH M.P., ZEMMOUR J., WARD F.E.,
RA WILLIAMS R.C., PARHAM P.;
RT "serologic cross-reactivities poorly reflect allelic relationships in
RT the HLA-B12 and HLA-B21 groups. Dominant epitopes of the alpha 2
RT helix.";
RL J. Immunol. 149:3563-3568(1992).
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC THE IMMUNE SYSTEM.
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC MICROGLOBULIN).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M24037; AAA02950.1; -.
DR HSSP; P30491; 1A1M.
DR MIM; 142830; -.
DR PROSITE; PS00290; IG_MHC; 1.
DR PFAM; PF00047; Ig; 1.
DR PFAM; PF00129; MHC_I; 1.
KW MHC I; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 24
FT CHAIN 25 362
FT FT HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
FT B-49(B-21) B*4901 ALPHA CHAIN.
FT FT EXTRACELLULAR ALPHA-1.
FT DOMAIN 25 114
FT DOMAIN 115 206
FT DOMAIN 207 298
FT DOMAIN 299 309
FT DOMAIN 309 333
FT TRANSMEM 310 333
FT DOMAIN 334 352
FT CARBOHYD 110 110
FT DISULFID 125 188
FT DISULFID 227 283
FT SEQUENCE 362 AA; 40581 MW; E996F82F CRC32;

```

```

Query Match 88.0%; Score 44; DB 1; Length 362;
Best Local Similarity 90.0%; Pred. No. 0.076;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 RENLTRY 10
Db 99 RENLTRY 108

RESULT 11
ID 1B49_HUMAN STANDARD; PRT; 362 AA.
AC P18464.
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-51(B-5) B*5101 ALPHA CHAIN
DE PRECURSOR.
GN HLA-B OR HLAB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE; 90207291.
RA ENNIS P.D., ZEMMOUR J., SALTER R.D., PARHAM P.;
RT "Rapid cloning of HLA-A,B cDNA by using the polymerase chain
RT reaction: frequency and nature of errors produced in amplification.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:2833-2837(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE; 89080265.

```

```

RA HAYASHI H., ENNIS P.D., ARIGA H., SALTER R.D., PARHAM P., KANO K.,
RA TAKIGUCHI M.;
RT "HLA-B51 and HLA-Bw52 differ by only two amino acids which are in the
RT helical region of the alpha 1 domain.";
RL J. Immunol. 142:306-311(1989).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE; 89233295.
RA POHLA H., KUON W., TABACZEWSKI P., DOERNER C., WEISS E.H.;
RT "Allelic variation in HLA-B and HLA-C sequences and the evolution of
RT the HLA-B alleles.";
RL Immunogenetics 29:297-307(1989).
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC THE IMMUNE SYSTEM.
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC MICROGLOBULIN).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M32319; AAA36232.1; -.
DR EMBL; M22792; AAA59620.1; ALT SEQ.
DR EMBL; M22786; AAA59620.1; JOINED.
DR EMBL; M22787; AAA59620.1; JOINED.
DR EMBL; M22788; AAA59620.1; JOINED.
DR EMBL; M22789; AAA59620.1; JOINED.
DR EMBL; M22790; AAA59620.1; JOINED.
DR EMBL; M22791; AAA59620.1; JOINED.
DR EMBL; L41086; AAA64513.1; -.
DR PIR; A30345; A30345.
DR PIR; A30548; A30548.
DR HSSP; P30491; 1A1M.
DR MIM; 142830; -.
DR PROSITE; PS00290; IG_MHC; 1.
DR PFAM; PF00047; Ig; 1.
DR PFAM; PF00129; MHC_I; 1.
KW MHC I; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 24
FT CHAIN 25 362
FT FT HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
FT B-51(B-5) B*5101 ALPHA CHAIN.
FT FT EXTRACELLULAR ALPHA-1.
FT DOMAIN 25 114
FT DOMAIN 115 206
FT DOMAIN 207 298
FT DOMAIN 299 308
FT TRANSMEM 309 332
FT DOMAIN 333 362
FT CARBOHYD 110 110
FT DISULFID 125 188
FT DISULFID 227 283
FT SEQUENCE 362 AA; 40566 MW; 4D846F30 CRC32;

```

```

Query Match 88.0%; Score 44; DB 1; Length 362;
Best Local Similarity 90.0%; Pred. No. 0.076;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 RENLTRY 10
Db 99 RENLTRY 108

RESULT 12
ID 1B52_HUMAN STANDARD; PRT; 362 AA.
AC P30489;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)

```

DT 01-APR-1993 (Rel. 25, Last annotation update)  
 DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B\*5104 ALPHA CHAIN  
 DE PRECURSOR.  
 GN HLA-B OR HLAB.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 9226955.  
 RA BELICH M.P., MADRIGAL J.A., HILDEBRAND W.H., ZEMMOUR J.,  
 RA WILLIAMS R.C., LUZ R., PETZL-ERLER M.L., PARHAM P.;  
 RL "Unusual HLA-B alleles in two tribes of Brazilian Indians."  
 RL Nature 357:326-329(1992).  
 CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO  
 CC THE IMMUNE SYSTEM.  
 CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
 CC MICROGLOBULIN).  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; Z15143; CAA78849.1; -  
 CC HSP; P30491; IALM.  
 CC MIN; 142830; -  
 CC PROSITE; PS00290; IG\_MHC; 1.  
 CC PFAM; PF00047; Ig; 1.  
 CC PFAM; PF00129; MHC.I; 1.  
 CC MHC I; Transmembrane; Glycoprotein; Signal.  
 CC SIGNAL 1 24  
 CC CHAIN 25 362 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,  
 CC B\*51(B-5) B\*5104 ALPHA CHAIN.  
 CC DOMAIN 25 114 EXTRACELLULAR ALPHA-1.  
 CC DOMAIN 115 206 EXTRACELLULAR ALPHA-2.  
 CC DOMAIN 207 298 EXTRACELLULAR ALPHA-3.  
 CC DOMAIN 299 308 CONNECTING PEPTIDE.  
 CC TRANSMEM 309 332  
 CC DOMAIN 333 362 CYTOPLASMIC TAIL.  
 CC CARBOHYD 110 110 BY SIMILARITY.  
 CC DISULFID 125 188 BY SIMILARITY.  
 CC DISULFID 227 283 BY SIMILARITY.  
 CC SEQUENCE 362 AA; 40560 MW; F22F08AB CRC32;  
 SQ  
 Query Match 88.0%; Score 44; DB 1; Length 362;  
 Best Local Similarity 90.0%; Pred. No. 0.076;  
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 RENLRTALRY 10  
 Db 99 RENLRTALRY 108  
 RESULT 13  
 ID 1B53\_HUMAN STANDARD; PRT; 362 AA.  
 AC P30490;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 01-APR-1993 (Rel. 25, Last annotation update)  
 DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, BW-52(B-5) B\*5201 ALPHA CHAIN  
 DE PRECURSOR.  
 GN HLA-B OR HLAB.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.

RX MEDLINE: 89080265.  
 RA HAYASHI H., ENNIS P.D., ARIGA H., SALTER R.D., PARHAM P., KANO K.,  
 RA TAKIGUCHI M.;  
 RT "HLA-B51 and HLA-Bw52 differ by only two amino acids which are in the  
 RT helical region of the alpha 1 domain."  
 RL J. Immunol. 142:306-311(1989).  
 CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO  
 CC THE IMMUNE SYSTEM.  
 CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
 CC MICROGLOBULIN).  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; M22799; AAA59645.1; ALT\_SEQ.  
 CC EMBL; M22793; AAA59645.1; JOINED.  
 CC EMBL; M22794; AAA59645.1; JOINED.  
 CC EMBL; M22795; AAA59645.1; JOINED.  
 CC EMBL; M22796; AAA59645.1; JOINED.  
 CC EMBL; M22797; AAA59645.1; JOINED.  
 CC EMBL; M22798; AAA59645.1; JOINED.  
 CC PIR; B30345; B30345.  
 CC PIR; B30548; B30548.  
 CC HSP; P30491; IALM.  
 CC MIN; 142830; -  
 CC PROSITE; PS00290; IG\_MHC; 1.  
 CC PFAM; PF00047; Ig; 1.  
 CC PFAM; PF00129; MHC.I; 1.  
 CC MHC I; Transmembrane; Glycoprotein; Signal.  
 CC SIGNAL 1 24  
 CC CHAIN 25 362 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,  
 CC BW-52(B-5) B\*5201 ALPHA CHAIN.  
 CC DOMAIN 25 114 EXTRACELLULAR ALPHA-1.  
 CC DOMAIN 115 206 EXTRACELLULAR ALPHA-2.  
 CC DOMAIN 207 298 EXTRACELLULAR ALPHA-3.  
 CC DOMAIN 299 308 CONNECTING PEPTIDE.  
 CC TRANSMEM 309 332  
 CC DOMAIN 333 362 CYTOPLASMIC TAIL.  
 CC CARBOHYD 110 110 BY SIMILARITY.  
 CC DISULFID 125 188 BY SIMILARITY.  
 CC DISULFID 227 283 BY SIMILARITY.  
 CC SEQUENCE 362 AA; 40521 MW; 3B436FE8 CRC32;  
 SQ  
 Query Match 88.0%; Score 44; DB 1; Length 362;  
 Best Local Similarity 90.0%; Pred. No. 0.076;  
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 RENLRTALRY 10  
 Db 99 RENLRTALRY 108  
 RESULT 14  
 ID 1B54\_HUMAN STANDARD; PRT; 362 AA.  
 AC P30491;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, BW-53 B\*5301 ALPHA CHAIN  
 DE PRECURSOR.  
 GN HLA-B OR HLAB.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.

RX MEDLINE: 91033941.  
 RA HAYASHI H., Ooba T., NAKAYAMA S., SEKIMATA M., KANO K.,  
 RA TAKIGUCHI M.;  
 RT "Allotopicities between HLA-Bw53 and HLA-B35 are generated by  
 RT substitution of the residues associated with HLA-Bw4/Bw6 public  
 RT epitopes.";  
 RL Immunogenetics 32:195-199(1990).  
 RN [2]  
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 25-302.  
 RX MEDLINE: 96209672.  
 RA SMITH K.J., REID S.W., HARLOS K., MCMICHAEL A.J., STUART D.I.,  
 RA BELL J.I., JONES E.Y.;  
 RT "Bound water structure and polymorphic amino acids act together to  
 RT allow the binding of different peptides to MHC class I HLA-B53.";  
 RL Immunity 4:215-228(1996).  
 CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO  
 CC THE IMMUNE SYSTEM.  
 CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
 CC MICROGLOBULIN).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL: M58636; AAA36228.1;  
 DR PIR: A45834; A45834.  
 DR PDB: 1A1M; 08-APR-98.  
 DR PDB: 1A1O; 08-APR-98.  
 DR MIM: 142830;  
 DR PROSITE: PS00290; IG\_MHC; 1.  
 DR PFAM: PF00047; ig; 1.  
 DR PFAM: PF00129; MHC\_I; 1.  
 KW MHC I; Transmembrane; Glycoprotein; Signal; 3D-structure.  
 FT SIGNAL 1 24  
 FT CHAIN 25 362  
 FT HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,  
 FT BW-53 B\*5301 ALPHA CHAIN.  
 FT DOMAIN 25 114  
 FT DOMAIN 115 206  
 FT DOMAIN 207 298  
 FT DOMAIN 299 308  
 FT TRANSMEM 309 332  
 FT DOMAIN 333 362  
 FT CARBOHYD 110 110  
 FT DISULFID 125 188  
 FT DISULFID 227 283  
 SQ SEQUENCE 362 AA; 40495 MW; 2BDC746E CRC32;

Query Match 88.0%; Score 44; DB 1; Length 362;  
 Best Local Similarity 90.0%; Pred. No. 0.076;  
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 RENLRTALRY 10  
 Db 99 RENLRTALRY 108

RESULT 15  
 ID 1B60\_HUMAN STANDARD; PRT; 362 AA.  
 AC P18465;  
 DT 01-NOV-1990 (Rel. 16, Created)  
 DT 01-NOV-1990 (Rel. 16, Last sequence update)  
 DT 01-APR-1993 (Rel. 25, Last annotation update)  
 DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-57(B-17) B\*5701 ALPHA  
 DE CHAIN PRECURSOR (BW57.1).  
 GN HLA-B OR HLAB  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 90207291.  
 RA ENNIS P.D., ZEMMOUR J., SALTER R.D., PARHAM P.;  
 RT "Rapid cloning of HLA-A,B CDNA by using the polymerase chain  
 RT reaction: frequency and nature of errors produced in amplification.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 87:2833-2837(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 91067476.  
 RA ISAMAT M., GRIDLESTONE J., MILSTEIN C.;  
 RT "Nucleotide sequence of an HLA-Bw57 gene.";  
 RL Nucleic Acids Res. 18:6702-6702(1990).  
 CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO  
 CC THE IMMUNE SYSTEM.  
 CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
 CC MICROGLOBULIN).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL: M32318; AAA36231.1;  
 DR EMBL: X55711; CAA39244.1;  
 DR PIR: S12622; S12622.  
 DR PIR: D35997; D35997.  
 DR HSSP: P30491; 1A1M.  
 DR MIM: 142830;  
 DR PROSITE: PS00290; IG\_MHC; 1.  
 DR PFAM: PF00047; ig; 1.  
 DR PFAM: PF00129; MHC\_I; 1.  
 KW MHC I; Transmembrane; Glycoprotein; Signal.  
 FT SIGNAL 1 24  
 FT CHAIN 25 362  
 FT HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,  
 FT B-57(B-17) B\*5701 ALPHA CHAIN.  
 FT DOMAIN 25 114  
 FT DOMAIN 115 206  
 FT DOMAIN 207 298  
 FT DOMAIN 299 308  
 FT TRANSMEM 309 332  
 FT DOMAIN 333 362  
 FT CARBOHYD 110 110  
 FT DISULFID 125 188  
 FT DISULFID 227 283  
 SQ SEQUENCE 362 AA; 40224 MW; D91DF8DD CRC32;

Query Match 88.0%; Score 44; DB 1; Length 362;  
 Best Local Similarity 90.0%; Pred. No. 0.076;  
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 RENLRTALRY 10  
 Db 99 RENLRTALRY 108

Search completed: February 8, 2000, 01:25:58  
 Job time: 1558 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2000, 19:16:13 ; Search time 176.54 Seconds  
(without alignments)  
3.927 Million cell updates/sec

Title: US-08-653-294-25  
Perfect score: 50  
Sequence: 1 RENLRTALRY 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 225878 seqs, 59334122 residues  
Total number of hits satisfying chosen parameters: 225878

Minimum DB seq length: 0  
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database : SPTREMBL\_12.\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50	100.0	89	7 019674	019674 homo sapien
2	50	100.0	89	7 019565	019565 homo sapien
3	50	100.0	131	7 097998	097998 homo sapien
4	50	100.0	131	7 097999	097999 homo sapien
5	50	100.0	181	7 019779	019779 homo sapien
6	50	100.0	181	7 030197	030197 homo sapien
7	50	100.0	181	7 019669	019669 homo sapien
8	50	100.0	181	7 078028	078028 homo sapien
9	50	100.0	355	7 029853	029853 homo sapien
10	50	100.0	361	7 09XRVO	09XRVO pongo pygma
11	50	100.0	362	7 029637	029637 homo sapien
12	50	100.0	362	7 029935	029935 homo sapien
13	50	100.0	362	7 079524	079524 homo sapien
14	50	100.0	362	7 029850	029850 homo sapien
15	50	100.0	362	7 029661	029661 homo sapien
16	50	100.0	362	7 078180	078180 homo sapien
17	50	100.0	362	7 029933	029933 homo sapien
18	45	90.0	89	4 095956	095956 homo sapien
19	45	90.0	89	7 077959	077959 homo sapien
20	45	90.0	137	7 095533	095533 pan troglod

21	45	90.0	181	7 077934	077934 homo sapien
22	45	90.0	330	7 019356	019356 macaca mula
23	45	90.0	331	7 002944	002944 macaca mula
24	45	90.0	331	7 002945	002945 macaca mula
25	45	90.0	345	7 095459	095459 rattus norv
26	45	90.0	345	7 078088	078088 rattus norv
27	45	90.0	348	7 046875	046875 rattus norv
28	45	90.0	362	7 029938	029938 homo sapien
29	44	88.0	89	7 019569	019569 homo sapien
30	44	88.0	90	7 046693	046693 pan troglod
31	44	88.0	90	7 046694	046694 pan troglod
32	44	88.0	90	7 046695	046695 pan troglod
33	44	88.0	90	7 046696	046696 pan paniscu
34	44	88.0	90	7 046697	046697 gorilla gor
35	44	88.0	133	7 019189	019189 homo sapien
36	44	88.0	138	7 078209	078209 homo sapien
37	44	88.0	172	7 019770	019770 homo sapien
38	44	88.0	172	7 019774	019774 homo sapien
39	44	88.0	172	7 019775	019775 homo sapien
40	44	88.0	172	7 019780	019780 homo sapien
41	44	88.0	172	7 095364	095364 homo sapien
42	44	88.0	172	7 019771	019771 homo sapien
43	44	88.0	172	7 019772	019772 homo sapien
44	44	88.0	172	7 019773	019773 homo sapien
45	44	88.0	175	7 029694	029694 homo sapien

ALIGNMENTS

RESULT 1  
ID 019674 PRELIMINARY; PRT: 89 AA.  
AC 019674;  
DT 01-JAN-1998 (TREMBLrel. 05, Created)  
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
DE 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
DE HLA-B\*13 (FRAGMENT).  
GN HLA-B\*13.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA MAERTENS R., DE CANCK I.;  
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; Y12378; CAA73021.1; -  
DR PFAM; PF00129; MHC\_I; 1.  
KW MHC.  
FT NON\_TER 1 1  
FT NON\_TER 89 89  
SQ SEQUENCE 89 AA; 10581 MW; 9AC7631C CRC32;  
  
Query Match 100.0%; Score 50; DB 7; Length 89;  
Best Local Similarity 100.0%; Pred. No. 0.009;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 RENLRTALRY 10  
Db 74 RENLRTALRY 83  
  
RESULT 2  
ID 019565 PRELIMINARY; PRT: 89 AA.  
AC 019565;  
DT 01-JAN-1998 (TREMBLrel. 05, Created)  
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)  
DE MHC CLASS I ANTIGEN (FRAGMENT).  
GN HLA-B.  
OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA CAO K., BURDETT L., ZHANG G., FERNANDEZ-VINA M.;  
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF017316; AAB70282.2; -.  
 KW MHC.  
 FT NON\_TER 1 1  
 FT NON\_TER 89 89  
 SQ SEQUENCE 89 AA; 10581 MW; 9AC7631C CRC32;

Query Match 100.0%; Score 50; DB 7; Length 89;  
 Best Local Similarity 100.0%; Pred. No. 0.009;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RENLRTALRY 10  
 Db 74 RENLRTALRY 83

## RESULT 3

O97998  
 ID O97998 PRELIMINARY; PRT; 131 AA.  
 AC O97998;  
 DT 01-MAY-1999 (TRENBLrel. 10, Created)  
 DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TRENBLrel. 10, Last annotation update)  
 DE MHC CLASS I ANTIGEN (FRAGMENT).  
 GN HLA-B.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA HOLDSWORTH R.;  
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF035648; AAD02035.1; -.  
 KW MHC.  
 FT NON\_TER 1 1  
 FT NON\_TER 131 131  
 SQ SEQUENCE 131 AA; 15134 MW; 931C8D27 CRC32;

Query Match 100.0%; Score 50; DB 7; Length 131;  
 Best Local Similarity 100.0%; Pred. No. 0.013;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RENLRTALRY 10  
 Db 40 RENLRTALRY 49

## RESULT 4

O97999  
 ID O97999 PRELIMINARY; PRT; 131 AA.  
 AC O97999;  
 DT 01-MAY-1999 (TRENBLrel. 10, Created)  
 DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TRENBLrel. 10, Last annotation update)  
 DE MHC CLASS I ANTIGEN (FRAGMENT).  
 GN HLA-B.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA HOLDSWORTH R.;  
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF035649; AAD02036.1; -.  
 KW MHC.  
 FT NON\_TER 1 1

FT NON\_TER 131 131  
 SQ SEQUENCE 131 AA; 15057 MW; 32C865A3 CRC32;  
 Query Match 100.0%; Score 50; DB 7; Length 131;  
 Best Local Similarity 100.0%; Pred. No. 0.013;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RENLRTALRY 10  
 Db 40 RENLRTALRY 49

## RESULT 5

O19779  
 ID O19779 PRELIMINARY; PRT; 181 AA.  
 AC O19779;  
 DT 01-JAN-1998 (TRENBLrel. 05, Created)  
 DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)  
 DT 01-NOV-1999 (TRENBLrel. 12, Last annotation update)  
 DE MHC CLASS I ANTIGEN (FRAGMENT).  
 GN HLA-B.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STEINER N.K., NG J., BUSH J., HARTZMAN R.J., JOHNSON-DOW L.,  
 RA HURLEY C.K.;  
 RL Hum. Immunol. 56:0-0(1997).  
 DR EMBL; U90241; AAB82305.1; -.  
 DR HSSP; P10318; 1ROG.  
 DR PFAM; PF00129; MHC\_I; 1.  
 KW MHC.  
 FT NON\_TER 1 1  
 FT NON\_TER 181 181  
 SQ SEQUENCE 181 AA; 21095 MW; 97EC2597 CRC32;

Query Match 100.0%; Score 50; DB 7; Length 181;  
 Best Local Similarity 100.0%; Pred. No. 0.019;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RENLRTALRY 10  
 Db 74 RENLRTALRY 83

## RESULT 6

O30197  
 ID O30197 PRELIMINARY; PRT; 181 AA.  
 AC O30197;  
 DT 01-NOV-1996 (TRENBLrel. 01, Created)  
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
 DT 01-NOV-1999 (TRENBLrel. 12, Last annotation update)  
 DE MHC CLASS I ANTIGEN (FRAGMENT).  
 GN HLA-B.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA HURLEY C.K., HOYER R.J.;  
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U63560; AAB05925.1; -.  
 DR EMBL; U63559; AAB05925.1; JOINED.  
 DR HSSP; P10318; 1ROG.  
 DR PFAM; PF00129; MHC\_I; 1.  
 KW MHC.  
 FT NON\_TER 1 1  
 FT NON\_TER 181 181  
 SQ SEQUENCE 181 AA; 20997 MW; 7DFABE37 CRC32;



Query Match 100.0%; Score 50; DB 7; Length 181;  
 Best Local Similarity 100.0%; Pred. No. 0.019;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RENLRTALRY 10  
 Db 74 RENLRTALRY 83  
 |||||

RESULT 7  
 ID O19669 PRELIMINARY; PRT; 181 AA.  
 AC O19669;  
 DT 01-JAN-1998 (TREMBlrel. 05, Created)  
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
 DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)  
 DE MHC CLASS I ANTIGEN SHCHA (FRAGMENT).  
 GN HLA-B.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA HURLEY C.K., STEINER N.K.;  
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U58470; AAB02607.1; -.  
 DR EMBL; U58469; AAB02607.1; JOINED.  
 DR HSP; P10318; IROG.  
 DR PFAM; PF00129; MHC\_I; 1.  
 KW MHC.  
 FT NON\_TER 1 181  
 FT NON\_TER 181  
 SQ SEQUENCE 181 AA; 21028 MW; EC872642 CRC32;

Query Match 100.0%; Score 50; DB 7; Length 181;  
 Best Local Similarity 100.0%; Pred. No. 0.019;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RENLRTALRY 10  
 Db 74 RENLRTALRY 83  
 |||||

RESULT 8  
 ID O78028 PRELIMINARY; PRT; 181 AA.  
 AC O78028;  
 DT 01-NOV-1998 (TREMBlrel. 08, Created)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)  
 DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)  
 DE HLA-B\*15MD. (FRAGMENT).  
 GN HLA-B\*15MD.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=BLOOD;  
 RA GAO X., MATHESON B.;  
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U58316; AAB87723.1; -.  
 DR EMBL; U58315; AAB87723.1; JOINED.  
 DR HSP; P10318; IROG.  
 DR PFAM; PF00129; MHC\_I; 1.  
 KW MHC.  
 FT NON\_TER 1 181  
 FT NON\_TER 181  
 SQ SEQUENCE 181 AA; 21029 MW; C35A18BE CRC32;

Query Match 100.0%; Score 50; DB 7; Length 181;  
 Best Local Similarity 100.0%; Pred. No. 0.019;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RENLRTALRY 10  
 Db 74 RENLRTALRY 83  
 |||||

RESULT 9  
 ID Q29853 PRELIMINARY; PRT; 355 AA.  
 AC Q29853;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)  
 DE HLA-B ALPHA-CHAIN (FRAGMENT).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 95176328.  
 RA GAUCHAT-FEISS D., BREUR-VRISENDORP B.S., RUFER N., JEANNET M.,  
 RA ROOSNEK E., TIERCY J.M.;  
 RT "Sequencing of a novel functional HLA-B\*44 subtype differing in two  
 RT residues in the alpha 2 domain.";  
 RL Tissue Antigens 44:261-264(1994).  
 DR EMBL; X75953; CAA53566.1; -.  
 DR HSP; P30491; IALM.  
 DR PROSITE; PS00290; IG\_MHC; 1.  
 DR PFAM; PF00047; Ig; 1.  
 DR PFAM; PF00129; MHC\_I; 1.  
 KW MHC.  
 FT NON\_TER 355 355  
 FT NON\_TER 355  
 SQ SEQUENCE 355 AA; 39845 MW; EF81934E CRC32;

Query Match 100.0%; Score 50; DB 7; Length 355;  
 Best Local Similarity 100.0%; Pred. No. 0.037;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RENLRTALRY 10  
 Db 99 RENLRTALRY 108  
 |||||

RESULT 10  
 ID Q9XRY0 PRELIMINARY; PRT; 361 AA.  
 AC Q9XRY0;  
 DT 01-NOV-1999 (TREMBlrel. 12, Created)  
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
 DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)  
 DE MHC CLASS I ANTIGEN (FRAGMENT).  
 GN POPY B.  
 OS Pongo pygmaeus (Orangutan).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Pongo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA ADAMS E.J., THOMSON G., PARHAM P.;  
 RT "Evidence for an HLA-C-like locus in the orangutan Pongo pygmaeus.";  
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF118895; AAD28435.1; -.  
 DR PROSITE; PS00290; IG\_MHC; 1.  
 DR NON\_TER 1 361  
 FT NON\_TER 1  
 SQ SEQUENCE 361 AA; 40466 MW; CB177B4F CRC32;

Query Match 100.0%; Score 50; DB 7; Length 361;  
 Best Local Similarity 100.0%; Pred. No. 0.038;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RENLRTALRY 10  
| | | | |  
Db 98 RENLRTALRY 107

RESULT 11  
Q29637 PRELIMINARY; PRT; 362 AA.  
AC Q29637;  
DT 01-NOV-1996 (TRENBLrel. 01, Created)  
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
DT 01-NOV-1999 (TRENBLrel. 12, Last annotation update)  
DE MHC CLASS I ANTIGEN.  
GN HLA-B.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA DOMENA J.D., PARHAM P.;  
RL Submitted (DEC-1993) to the EMBL/GenBank/DBJ databases.  
DR EMBL: U04244; AAA87397.1; -.  
DR HSSP: P30460; IAGE.  
DR PROSITE: PS00290; IG\_MHC; 1.  
DR PFAM: PF00047; ig; 1.  
DR PFAM: PF00129; MHC\_I; 1.  
KW MHC.  
SQ SEQUENCE 362 AA; 40359 MW; EEAB95D7 CRC32;

Query Match 100.0%; Score 50; DB 7; Length 362;  
Best Local Similarity 100.0%; Pred. No. 0.038;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RENLRTALRY 10  
| | | | |  
Db 99 RENLRTALRY 108

RESULT 12  
Q29935 PRELIMINARY; PRT; 362 AA.  
AC Q29935;  
DT 01-NOV-1996 (TRENBLrel. 01, Created)  
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
DT 01-NOV-1999 (TRENBLrel. 12, Last annotation update)  
DE MHC CLASS I HLA-B\*3802.  
GN HLA-B.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA LITTLE A.M., DOMENA J.D., HILDEBRAND W.H., PARHAM P.;  
RL Submitted (FEB-1994) to the EMBL/GenBank/DBJ databases.  
DR EMBL: L22028; AAA59618.1; -.  
DR HSSP: P30491; IALM.  
DR PROSITE: PS00290; IG\_MHC; 1.  
DR PFAM: PF00047; ig; 1.  
DR PFAM: PF00129; MHC\_I; 1.  
KW MHC.  
SQ SEQUENCE 362 AA; 40404 MW; 83A8399C CRC32;

Query Match 100.0%; Score 50; DB 7; Length 362;  
Best Local Similarity 100.0%; Pred. No. 0.038;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RENLRTALRY 10  
| | | | |  
Db 99 RENLRTALRY 108

RESULT 13  
P79524 PRELIMINARY; PRT; 362 AA.  
AC P79524;  
DT 01-MAY-1997 (TRENBLrel. 03, Created)  
DT 01-MAY-1997 (TRENBLrel. 03, Last sequence update)  
DT 01-NOV-1999 (TRENBLrel. 12, Last annotation update)  
DE MHC CLASS I HLA-B\*44BO.  
GN HLA-B.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA ARNETT K.L., DARKE C., PARHAM P.;  
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
DR EMBL: U64801; AAB40632.1; -.  
DR HSSP: P30491; IALM.  
DR PROSITE: PS00290; IG\_MHC; 1.  
DR PFAM: PF00047; ig; 1.  
DR PFAM: PF00129; MHC\_I; 1.  
KW MHC.  
SQ SEQUENCE 362 AA; 40396 MW; AC0FAF3C CRC32;

Query Match 100.0%; Score 50; DB 7; Length 362;  
Best Local Similarity 100.0%; Pred. No. 0.038;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RENLRTALRY 10  
| | | | |  
Db 99 RENLRTALRY 108

RESULT 14  
Q29850 PRELIMINARY; PRT; 362 AA.  
AC Q29850;  
DT 01-NOV-1996 (TRENBLrel. 01, Created)  
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
DT 01-NOV-1999 (TRENBLrel. 12, Last annotation update)  
DE HLA-B ALPHA CHAIN B\*4407 PRECURSOR.  
GN HLA-B\*4407.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 97004417.  
RA VILCHES C., SANZ L., DE PABLO R., MORENO M.E., PUENTE S., KREISLER M.;  
RT "Molecular characterization of the new alleles HLA-B\*8101 and  
RT B\*4407".  
RL Tissue Antigens 47:139-142(1996).  
DR EMBL: X90391; CAA62036.1; -.  
DR HSSP: P30491; IALM.  
DR PROSITE: PS00290; IG\_MHC; 1.  
DR PFAM: PF00047; ig; 1.  
DR PFAM: PF00129; MHC\_I; 1.  
KW Signal; MHC.  
FT SIGNAL 1 24 POTENTIAL.  
FT CHAIN 25 362 HLA-B ALPHA CHAIN B\*4407.  
SQ SEQUENCE 362 AA; 40449 MW; 07B5245D CRC32;

Query Match 100.0%; Score 50; DB 7; Length 362;  
Best Local Similarity 100.0%; Pred. No. 0.038;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RENLRTALRY 10  
| | | | |  
Db 99 RENLRTALRY 108

```

RESULT 15
Q29661 ID Q29661 PRELIMINARY; PRT; 362 AA.
AC Q29661;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
DE MHC ANTIGEN HLA-B*1303 PRECURSOR.
GN HLA-B*1303.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Homidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BLOOD;
RX MEDLINE; 96232969.
RA BALAS A., GARCIA-SANCHEZ F., VICARIO J.L.;
RT "HLA-B*1303: a new example of poor correlation between serology and
RT structure.";
RL Hum. Immunol. 45:32-36(1996).
DR EMBL; U14943; AAB06829.1; -.
DR HSP; P30491; IALM.
DR PROSITE; PS00290; IG_MHC; 1.
DR PFAM; PF00047; Ig; 1.
DR PFAM; PF00129; MHC_I; 1.
KW Signal; MHC.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 362 MHC ANTIGEN HLA-B*1303.
SQ SEQUENCE 362 AA; 40501 MW; 9277DEFB CRC32;

```

```

Query Match 100.0%; Score 50; DB 7; Length 362;
Best Local Similarity 100.0%; Pred. No. 0.038;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Oy 1 RENTLALRY 10
Db 99 RENTLALRY 108

```

```

Search completed: February 8, 2000, 19:16:13
Job time: 21500 sec

```

**THIS PAGE BLANK (USPTO)**

OM of: US-08-653-294-25 to: GenEmbl:\* out\_format : pfs  
 Date: Feb 8, 2000 10:22 PM  
 About: Results were produced by the GenCore software, version 4.5,  
 Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:  
 -MODEL=framet p2n.model -DEV=xlp  
 -C/cnrl\_1/USPTO\_spool/US08653294/runat\_04022000\_160701\_15779/app\_query.fasta.2  
 -DB=GenEmbl -QFMT=fastap -SUFFIX=rg -GAPOP=12.000 -GAPEXT=4.000  
 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -OGAPOP=4.500  
 -CGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000  
 -DELEXT=7.000 -XGAPOP=10.000 -XGAPEXT=0.500 -DELOP=6.000  
 -DELEXT=7.000 -START=1 -MATRIX=blosum62 -TRANS=human40.cdi  
 -LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -ALIGN=15 -MODE=LOCAL  
 -OUTFMT=pfs -NORM=ext -MINLEN=0 -MAXLEN=1000000 -USER=US08653294  
 -NCPU=6 -ICPU=3 -NO\_XLPXY -WAIT -THREADS=1

Search information block:  
 Query: US-08-653-294-25  
 Query length: 10  
 Database: GenEmbl:\*  
 Database sequences: 821193  
 Database length: -1518192014  
 Search time (sec): 10176.920000

score_list:	Sequence	Strd Orig	Zscore	Escore	Len	Documentation
gb_pr1:HSB44HLA2	+	50.00	193.52	0.0093	163	X78449 H.sapiens HLA-B gene, ex
gb_pr1:HSB44HLA2	+	50.00	192.33	0.0109	187	X78426 H.sapiens HLA-B gene, ex
gb_pr1:HSB44HLA2	+	50.00	189.15	0.0164	270	X18648 Homo sapiens HLA-B gene
gb_pr1:HSB44HLA2	+	50.00	189.15	0.0164	270	AJ132859 Homo sapiens HLA-B gene
gb_pr1:HSB44HLA2	+	50.00	189.15	0.0164	270	X13378 H.sapiens HLA-B*13 gene
gb_pr1:HSB44HLA2	+	50.00	189.15	0.0164	270	AJ249724 Homo sapiens partial H
gb_pr1:HSB44HLA2	+	50.00	189.15	0.0164	270	U58469 Human MHC class I antige
gb_pr1:HSB44HLA2	+	50.00	189.15	0.0164	270	U58315 Human HLA-B*15MD gene, e
gb_pr1:HSB44HLA2	+	50.00	189.15	0.0164	270	U90240 Human MHC class I antige
gb_pr1:HSB44HLA2	+	50.00	189.15	0.0164	270	AF017316 Homo sapiens MHC class
gb_pr1:HSB44HLA2	+	50.00	189.15	0.0164	270	AF181857 Homo sapiens HLA-B MHC
gb_pr1:HSB44HLA2	+	50.00	189.15	0.0164	270	U63559 Human MHC class I antige
gb_pr1:HSB44HLA2	+	50.00	185.86	0.0250	395	AF035648 Homo sapiens MHC class
gb_pr1:HSB44HLA2	+	50.00	185.86	0.0250	395	AF035649 Homo sapiens MHC class
gb_pr1:HSB44HLA2	+	50.00	185.66	0.0256	404	L31798 Homo sapiens HLA-B*44 va
gb_pr1:HSB44HLA2	+	50.00	183.29	0.0347	531	AB032598 Homo sapiens HLA-B*1301
gb_pr1:HSB44HLA2	+	50.00	183.05	0.0357	546	AJ133267 Homo sapiens mRNA for
gb_pr1:HSB44HLA2	+	50.00	183.05	0.0357	546	L76935 Homo sapiens MHC class I
gb_pr1:HSB44HLA2	+	50.00	178.55	0.0637	918	AJ238702 Homo sapiens MHC class
gb_pr1:HSB44HLA2	+	50.00	178.28	0.0659	947	AF056981 Homo sapiens MHC class
gb_pr1:HSB44HLA2	+	50.00	177.28	0.0750	1064	X75953 H.sapiens HLA-B mRNA. 4
gb_pr1:HSB44HLA2	+	50.00	177.08	0.0769	1088	AF118895 Pongo pygmaeus MHC cl
gb_pr1:HSB44HLA2	+	50.00	177.07	0.0769	1089	X54366 H.sapiens HLA-B*4403 mR
gb_pr1:HSB44HLA2	+	50.00	177.07	0.0769	1089	X30391 H.sapiens mRNA for HLA-
gb_pr1:HSB44HLA2	+	50.00	177.07	0.0769	1089	D50290 Human mRNA for HLA-B*13
gb_pr1:HSB44HLA2	+	50.00	177.07	0.0769	1089	D50291 Human mRNA for HLA-B*13
gb_pr1:HSB44HLA2	+	50.00	177.07	0.0769	1089	L22028 Human MHC class I HLA-B
gb_pr1:HSB44HLA2	+	50.00	177.07	0.0769	1089	M24041 Human MHC class I HLA-B
gb_pr1:HSB44HLA2	+	50.00	177.07	0.0769	1089	M24038 Human MHC class I HLA-B
gb_pr1:HSB44HLA2	+	50.00	177.07	0.0769	1089	AJ131118 Homo sapiens mRNA for
gb_pr1:HSB44HLA2	+	50.00	177.07	0.0769	1089	U04244 Human MHC class I antige
gb_pr1:HSB44HLA2	+	50.00	177.07	0.0769	1089	U54801 Human MHC class I HLA-B
gb_pr1:HSB44HLA2	+	50.00	177.07	0.0769	1089	L42282 Homo sapiens (clone 18)
gb_pr1:HSB44HLA2	+	50.00	177.07	0.0769	1089	L42283 Homo sapiens major hist
gb_pr1:HSB44HLA2	+	50.00	177.07	0.0769	1089	U75533 Homo sapiens HLA class
gb_pr1:HSB44HLA2	+	50.00	177.04	0.0773	1093	M19757 Human MHC class I HLA-B
gb_pr1:HSB44HLA2	+	50.00	177.04	0.0773	1093	M24075 Human MHC class I HLA-B
gb_pr1:HSB44HLA2	+	50.00	175.53	0.0938	1301	U14943 Human MHC antigen (HLA-
gb_pr1:HSB44HLA2	+	50.00	145.60	4.36	41275	AC004182 Homo sapiens clone U
gb_pr1:HSB44HLA2	+	50.00	169.08	0.2146	270	AJ133471 Homo sapiens partial
gb_pr1:HSB44HLA2	+	50.00	169.08	0.2146	270	AF081275 Homo sapiens MHC class

gb\_pr3:HSB44HLA2 + 45.00 169.08 0.2146 270 ! AF0711767 Homo sapiens MHC cl  
 gb\_pr1:PTU05585 + 45.00 165.42 0.3432 412 ! U05585 Pan troglodytes class  
 gb\_ro:RNJ004889 + 45.00 157.39 0.9606 1041 ! AJ004889 Rattus norvegicus

seq\_name: gb\_pr1:HSB44HLA2

seq\_documentation\_block:  
 LOCUS HSB44HLA2 163 bp DNA PRI 16-DEC-1994  
 DEFINITION H.sapiens HLA-B gene, exon 2.  
 ACCESSION X78849  
 VERSION X78849.1 GI:502874  
 KEYWORDS heavy chain; MHC class I; MHC class I HLA-B\*44.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;  
 Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 163)  
 AUTHORS Yao,Z., Volgger,A., Scholz,S., Bonisch,J. and Albert,E.D.  
 TITLE Nucleotide sequence of a novel HLA-B\*44 subtype B\*4405  
 JOURNAL Immunogenetics 40 (4), 310 (1994)  
 MEDLINE 94364642  
 REFERENCE 2 (bases 1 to 163)  
 AUTHORS Albert,E.  
 TITLE Direct Submission  
 JOURNAL Submitted (11-APR-1994) E. Albert, Labor fuer Immunogenetik,  
 Kinderpoliklinik der LMU Muenchen, Pettenhoferstr. 8a, 80336  
 Muenchen, FRG

FEATURES  
 source Location/Qualifiers  
 1..163  
 /organism="Homo sapiens"  
 /isolate="WJG"  
 /db\_xref="taxon:9606"  
 /chromosome="6"  
 /gene="HLA-B"  
 /note="MHC"  
 1..163  
 /note="MHC Class I heavy chain HLA-B\*44 variant"  
 /number=2  
 BASE COUNT 45 a 50 c 54 g 14 t  
 ORIGIN

alignment\_scores:  
 Quality: 50.00 Length: 10  
 Ratio: 5.000 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
 US-08-653-294-25 x HSB44HLA2 ..  
 Align seg 1/1 to: HSB44HLA2 from: 1 to: 163

1 ArgGluAsnLeuArgThrAlaLeuArgTyr 10  
 |||||  
 115 CGAGAGAACCTGGCACCCTCCGCTAC 144

seq\_name: gb\_pr1:HSB44HLA2

seq\_documentation\_block:  
 LOCUS HSB44HLA2 187 bp DNA PRI 15-APR-1994  
 DEFINITION H.sapiens HLA-B gene, exon 2.  
 ACCESSION X78426  
 VERSION X78426.1 GI:473237  
 KEYWORDS heavy chain; HLA-B\*4404 gene; MHC class I associated.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;  
 Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 187)  
 AUTHORS Yao,Z. and Albert,E.  
 TITLE Identification of two major HLA-B\*44 subtypes as a novel B\*44  
 sequence: oligotyping  
 JOURNAL Unpublished

```

REFERENCE 2 (bases 1 to 187)
AUTHORS Yao, Z.
TITLE Direct Submission
JOURNAL Submitted (18-MAR-1994) Z. Yao, Immunogenetics Laboratory,
Kinderpoliklinik University of Munich, Pettenkaferstr 8a, 80336
Munich, FRG
REMARK revised by [3] MAT
REFERENCE 3 (bases 1 to 187)
AUTHORS Yao, Z.
TITLE Direct Submission
JOURNAL Submitted (15-APR-1994) Z. Yao, Immunogenetics Laboratory,
Kinderpoliklinik University of Munich, Pettenkaferstr 8a, 80336
Munich, FRG
COMMENT On Apr 17, 1994 this sequence version replaced gi:468539.
FEATURES
Source
Location/Qualifiers
1..187
/organism="Homo sapiens"
/isolate="BEB"
/db_xref="taxon:9606"
/chromosome="6"
/cell_line="BEB"
1..187
/gene="HLA-B44"
<1..187
/gene="HLA-B44"
/number=2
BASE COUNT 49 a 56 c 62 g 20 t
ORIGIN
alignment_scores:
Quality: 50.00 Length: 10
Ratio: 5.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-08-653-294-25 x HSHLAB2 ..
Align seg 1/1 to: HSHLAB2 from: 1 to: 187
1 ArgGluAsnLeuArgThrAlaLeuArgTyr 10
|||||
139 CGAGAGAACCTGCGACCGCTCCGCTAC 168
seq_name: gb_pr1:HSA18648
seq_documentation_block:
LOCUS HSA18648 270 bp DNA PRI 15-JAN-1999
DEFINITION Homo sapiens HLA-B gene, allele HLA-B*0802, exon 2.
ACCESSION Y18648
VERSION Y18648.1 GI:4160522
KEYWORDS HLA-B gene; HLA-B*0802 allele; human leucocyte antigen B; major
histocompatibility complex; MHC class I antigen.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
Guttridge, M.G.
1 (bases 1 to 270)
Direct Submission
Submitted (12-JAN-1999) M.G. Guttridge, Welsh Blood Service, Ely
Valley Road, Talbot Green, Pontyclun, CF72 9WB, UK
AUTHORS Guttridge, M.G.
TITLE Direct Submission
JOURNAL Submitted (12-JAN-1999) M.G. Guttridge, Welsh Blood Service, Ely
Valley Road, Talbot Green, Pontyclun, CF72 9WB, UK
FEATURES
Source
Location/Qualifiers
1..270
/organism="Homo sapiens"
/isolate="individual 22658"
/db_xref="taxon:9606"
/rearranged
/cell_type="leucocyte"
/tissue_type="blood"
/cell_line="22658"
<1..270
/gene="HLA-B"
exon

```

```

/number=2
/number=2
1..270
/gene="HLA-B"
<1..270
/gene="HLA-B"
/standard_name="MHC class I antigen"
/number=2
/number=2
/codon_start=3
/product="human leucocyte antigen B alpha chain"
/protein_id="CAA77241.1"
/db_xref="GI:4160523"
/translation="SHSMRYFDFTAMSPGCGEPFRFISGVYDDTQVRFDSDAASPRE
EPAPWIEGPEYWDNRNQTQIKNTQTDRENLRTALRYNQSEA"
BASE COUNT 60 a 88 c 87 g 35 t
ORIGIN
alignment_scores:
Quality: 50.00 Length: 10
Ratio: 5.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-08-653-294-25 x HSA18648 ..
Align seg 1/1 to: HSA18648 from: 1 to: 270
1 ArgGluAsnLeuArgThrAlaLeuArgTyr 10
|||||
222 CGAGAGAACCTGCGACCGCTCCGCTAC 251
seq_name: gb_pr1:HSA132659
seq_documentation_block:
LOCUS HSA132659 270 bp DNA PRI 02-FEB-1999
DEFINITION Homo sapiens HLA-B gene, B*4408 allele, exon 2 and joined CDS.
ACCESSION AJ132659
VERSION AJ132659.1 GI:4218239
KEYWORDS HLA-B gene; HLA-B*4408 allele; human leucocyte antigen B; major
histocompatibility complex; MHC class I antigen.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
Guttridge, M.G.
1 (bases 1 to 270)
Direct Submission
Submitted (01-FEB-1999) Guttridge M.G., Tissue Typing, Welsh Blood
Service, Ely Valley Road, Talbot Green, Pontyclun, Cardiff, CF72
9WB, UNITED KINGDOM
2 (bases 1 to 270)
Guttridge, M.G.
AUTHORS Guttridge, M.G.
TITLE Serology and confirmatory sequence of HLA-B*4408
JOURNAL Unpublished
FEATURES
Source
Location/Qualifiers
1..270
/organism="Homo sapiens"
/db_xref="taxon:9606"
/sex="Male"
/cell_type="lymphocytes"
/cell_line="ttid 30458"
1..270
/gene="HLA-B"
/number=2
/number=2
Join(<1..270,AJ132660.1:1..>276)
/gene="HLA-B"
Join(<1..270,AJ132660.1:1..>276)
/gene="HLA-B"
/standard_name="MHC class I antigen"
/number=2
/number=2
/codon_start=3

```

```

/label-hlaB_cds
/product="human leucocyte antigen B"
/protein_id="CAA10726.1"
/db_xref="GI:4218240"
/translation="SHSMRYFTAMSPRGEPFTITGVGVDDTLFVRFDSDAASPRM
APRAPTEQEGPYWDRETOISKNTQTYRENLTALRYNQSEAGSHIIQRMIGCDV
GPGRLRGYDDQAYDGKDYALNDELSSWTAADTAQAITQKWEAAARVAEQRAYLE
GLCVESLRYLENGKETLQRA"
BASE COUNT      57 a    91 c    85 g    37 t
ORIGIN

alignment_scores:
  Quality: 50.00      Length: 10
  Ratio: 5.000       Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-25 x HSA132659 ..
Align seg 1/1 to: HSA132659 from: 1 to: 270
1 ArgGluAsnLeuArgThrAlaLeuArgTyr 10
|||||
222 CGAGAGAACCTGCGCACCCTCGCTAC 251

seq_name: gb_pr1:HSLAB132

seq_documentation_block:
LOCUS      HSLAB132      270 bp      DNA      PRI      28-APR-1997
DEFINITION H.sapiens HLA-B*13 gene, variant exon 2.
ACCESSION Y12378
VERSION Y12378.1 GI:1934896
KEYWORDS exon 2; HLA-B*13.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 270)
AUTHORS Maertens,R. and De Canck,I.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 270)
AUTHORS De Canck,I.
TITLE Direct Submission
JOURNAL Submitted (07-APR-1997) I. De Canck, INNOGENETICS N.V.,
Industriepark Zwijnaarde 7 box 4, B-9052 Gent, BELGIUM

FEATURES
  source
    1..270
    /organism="Homo sapiens"
    /isolate="76002"
    /db_xref="taxon:9606"
    /chromosome="6"
    /map="p"
    <1..>270
    /gene="HLA-B*13"
    /number=2
    1..270
    /gene="HLA-B*13"
    <1..>270
    /codon_start=3
    /protein_id="CAA73021.1"
    /db_xref="GI:1934897"
    /db_xref="SPTREMBL:O19674"
    /translation="SHSMRYFTAMSPRGEPFTITGVGVDDTLFVRFDSDAASPRM
    APRAPTEQEGPYWDRETOISKNTQTYRENLTALRYNQSEA"
BASE COUNT      59 a    92 c    83 g    36 t
ORIGIN

alignment_scores:
  Quality: 50.00      Length: 10
  Ratio: 5.000       Gaps: 0

/label-hlaB_cds
/product="human leucocyte antigen B"
/protein_id="CAA10726.1"
/db_xref="GI:4218240"
/translation="SHSMRYFTAMSPRGEPFTITGVGVDDTLFVRFDSDAASPRM
APRAPTEQEGPYWDRETOISKNTQTYRENLTALRYNQSEAGSHIIQRMIGCDV
GPGRLRGYDDQAYDGKDYALNDELSSWTAADTAQAITQKWEAAARVAEQRAYLE
GLCVESLRYLENGKETLQRA"
BASE COUNT      57 a    91 c    85 g    37 t
ORIGIN

alignment_scores:
  Quality: 50.00      Length: 10
  Ratio: 5.000       Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-25 x HSLAB132 ..
Align seg 1/1 to: HSLAB132 from: 1 to: 270
1 ArgGluAsnLeuArgThrAlaLeuArgTyr 10
|||||
222 CGAGAGAACCTGCGCACCCTCGCTAC 251

seq_name: gb_pr2:HS249724

seq_documentation_block:
LOCUS      HS249724      270 bp      DNA      PRI      06-OCT-1999
DEFINITION Homo sapiens partial HLA-B gene for MHC class I antigen, HLA-B*44
variant, exon 2.
ACCESSION AJ249724
VERSION AJ249724.1 GI:5919229
KEYWORDS HLA-B gene; HLA-B*44 variant; human leucocyte antigen B; major
histocompatibility complex; MHC class I antigen.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 270)
AUTHORS Elsner,H.A., Schmitz,G. and Blasczyk,R.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 270)
AUTHORS Blasczyk,R.
TITLE Direct Submission
JOURNAL Submitted (17-SEP-1999) Blasczyk R., Department of Transfusion
Medicine, Hannover Medical School, Carl-Neuberg-Str. 1, Hannover,
30625, GERMANY

FEATURES
  Location/Qualifiers
    1..270
    /organism="Homo sapiens"
    /isolate="MH997336/Bt1268"
    /db_xref="taxon:9606"
    /chromosome="6"
    /map="6p21"
    1..270
    /gene="HLA-B"
    /note="HLA-B*44 variant"
    /number=2
    1..270
    /gene="HLA-B"
    <1..>270
    /gene="HLA-B"
    /standard_name="MHC class I antigen"
    /note="HLA-B*44 variant"
    /codon_start=3
    /label="HLA-B*44_CDS"
    /product="human leucocyte antigen B"
    /protein_id="CAB56342.2"
    /db_xref="GI:6015538"
    /translation="SHSMRYFTAMSPRGEPFTITGVGVDDTLFVRFDSDAASPRK
    EPRAPTEQEGPYWDRETOISKNTQTYRENLTALRYNQSEA"
BASE COUNT      60 a    89 c    85 g    36 t
ORIGIN

alignment_scores:
  Quality: 50.00      Length: 10
  Ratio: 5.000       Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-25 x HSA249724 ..
Align seg 1/1 to: HSA249724 from: 1 to: 270

```

```

1 ArgGluAsnLeuArgThrAlaLeuArgTyr 10
|||||
222 CGAGAGAACCTGCGCACCGCTCCGCTAC 251

seq_name: gb_pr2:HSSHCHA01

seq_documentation_block:
LOCUS HSSHCHA01 270 bp DNA 18-JUN-1996
DEFINITION Human MHC class I antigen SHCHA (HLA-B*4403 variant) gene, exon 2.
ACCESSION U58469
VERSION U58469.1 GI:1378136
KEYWORDS
SEGMENT
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
1 (bases 1 to 270)
AUTHORS Hurley,C.K. and Steiner,N.K.
TITLE Novel HLA-B alleles
JOURNAL Unpublished
REFERENCE
2 (bases 1 to 270)
AUTHORS Hurley,C.K. and Steiner,N.K.
TITLE Direct Submission
JOURNAL Submitted (16-MAY-1996) C.K. Hurley, Microbiology & Immunology,
Georgetown University, 3970 Reservoir Rd.NW, Washington, DC 20007,
USA
FEATURES
Source
1..270
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_type="PBL"
/haplotype="B*5101/B*4403 variant"
/chromosome="6"
/note="SHCHA-B*4403 variant, Direct sequence"
1..270
/gene="HLA-B"
/note="HLA-B*4403 variant"
/number=2
BASE COUNT 60 a 89 c 85 g 36 t
ORIGIN

exon
1..270
Quality: 50.00 Length: 10
Ratio: 5.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-25 x HSSHCHA01
..
Align seg 1/1 to: HSSHCHA01 from: 1 to: 270
1 ArgGluAsnLeuArgThrAlaLeuArgTyr 10
|||||
222 CGAGAGAACCTGCGCACCGCTCCGCTAC 251

seq_name: gb_pr3:HSHLABS1

seq_documentation_block:
LOCUS HSHLABS1 270 bp DNA 30-OCT-1997
DEFINITION Human MHC class I antigen HLA-B GN00155-B*38022 gene, exon 2.
ACCESSION U90240
VERSION U90240.1 GI:1905853
KEYWORDS
SEGMENT
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
1 (bases 1 to 270)
AUTHORS Steiner,N.K., Ng,J., Bush,J., Hartzman,R.J., Johnson-Dow,L. and
Hurley,C.K.
TITLE HLA-B Alleles Associated with the B15 Serologically Defined
Antigens
JOURNAL Hum. Immunol. 56 (1997) In press
REFERENCE
2 (bases 1 to 270)
AUTHORS Steiner,N.K. and Hurley,C.K.
TITLE Direct Submission
JOURNAL Submitted (20-FEB-1997) Microbiology & Immunology, Georgetown
University, 3970 Reservoir Rd. NW, Washington, DC 20007, USA
FEATURES
Source
1..270
Location/Qualifiers
/organism="Homo sapiens"
/isolate="GN00155"
/db_xref="taxon:9606"
1..270
/gene="HLA-B"
/number=2
BASE COUNT 60 a 88 c 85 g 37 t
ORIGIN

exon
1..270
Quality: 50.00 Length: 10
Ratio: 5.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-25 x HSSHCHA01
..
Align seg 1/1 to: HSSHCHA01 from: 1 to: 270
1 ArgGluAsnLeuArgThrAlaLeuArgTyr 10
|||||
222 CGAGAGAACCTGCGCACCGCTCCGCTAC 251

seq_name: gb_pr3:HSHLABS1

seq_documentation_block:
LOCUS HSHLABS1 270 bp DNA 30-OCT-1997
DEFINITION Human MHC class I antigen HLA-B GN00155-B*38022 gene, exon 2.
ACCESSION U90240
VERSION U90240.1 GI:1905853
KEYWORDS
SEGMENT
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
1 (bases 1 to 270)
AUTHORS Steiner,N.K., Ng,J., Bush,J., Hartzman,R.J., Johnson-Dow,L. and
Hurley,C.K.
TITLE HLA-B Alleles Associated with the B15 Serologically Defined
Antigens
JOURNAL Hum. Immunol. 56 (1997) In press
REFERENCE
2 (bases 1 to 270)
AUTHORS Steiner,N.K. and Hurley,C.K.
TITLE Direct Submission
JOURNAL Submitted (20-FEB-1997) Microbiology & Immunology, Georgetown
University, 3970 Reservoir Rd. NW, Washington, DC 20007, USA
FEATURES
Source
1..270
Location/Qualifiers
/organism="Homo sapiens"
/isolate="GN00155"
/db_xref="taxon:9606"
1..270
/gene="HLA-B"
/number=2
BASE COUNT 60 a 88 c 85 g 37 t
ORIGIN

exon
1..270
Quality: 50.00 Length: 10
Ratio: 5.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-25 x HSSHCHA01
..
Align seg 1/1 to: HSSHCHA01 from: 1 to: 270
1 ArgGluAsnLeuArgThrAlaLeuArgTyr 10
|||||
222 CGAGAGAACCTGCGCACCGCTCCGCTAC 251

seq_name: gb_pr3:HSHLABS1

seq_documentation_block:
LOCUS HSHLABS1 270 bp DNA 01-DEC-1997
DEFINITION Human HLA-B*15MD gene, exon 2.
ACCESSION U58315
VERSION U58315.1 GI:2654406
KEYWORDS
SEGMENT
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
1 (sites)
AUTHORS Gao,X. and Matheson,B.
TITLE A novel b15 variant found in oceanic populations

```

```

JOURNAL Unpublished
REFERENCE 2 (bases 1 to 270)
AUTHORS Gao, X.
TITLE Direct Submission
JOURNAL Submitted (15-MAY-1996) Xiaojiang Gao, The Australian National
University, Human Genetics Group, Acton, Canberra, Act, Australia,
2601
FEATURES
Source
1..270
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/map="6p21"
/cell_line="MD674"
/cell_type="lymphocyte"
/tissue_type="blood"
1..270
/gene="HLA-B*15MD"
/number=2
BASE COUNT 58 a 91 c 85 g 36 t
ORIGIN

exon
1..270
Quality: 50.00 Length: 10
Ratio: 5.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-25 x HSHLAB15M1
..
Align seg 1/1 to: HSHLAB15M1 from: 1 to: 270
1 ArgGluAsnLeuArgThrAlaLeuArgTyr 10
|||||
222 CGAGAGAACCTGCGCACCGCTCCGCTAC 251

seq_name: gb_pr3:HSHLABSH1

seq_documentation_block:
LOCUS HSHLABSH1 270 bp DNA 30-OCT-1997
DEFINITION Human MHC class I antigen HLA-B GN00155-B*38022 gene, exon 2.
ACCESSION U90240
VERSION U90240.1 GI:1905853
KEYWORDS
SEGMENT
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
1 (bases 1 to 270)
AUTHORS Steiner,N.K., Ng,J., Bush,J., Hartzman,R.J., Johnson-Dow,L. and
Hurley,C.K.
TITLE HLA-B Alleles Associated with the B15 Serologically Defined
Antigens
JOURNAL Hum. Immunol. 56 (1997) In press
REFERENCE
2 (bases 1 to 270)
AUTHORS Steiner,N.K. and Hurley,C.K.
TITLE Direct Submission
JOURNAL Submitted (20-FEB-1997) Microbiology & Immunology, Georgetown
University, 3970 Reservoir Rd. NW, Washington, DC 20007, USA
FEATURES
Source
1..270
Location/Qualifiers
/organism="Homo sapiens"
/isolate="GN00155"
/db_xref="taxon:9606"
1..270
/gene="HLA-B"
/number=2
BASE COUNT 60 a 88 c 85 g 37 t
ORIGIN

exon
1..270
Quality: 50.00 Length: 10
Ratio: 5.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-25 x HSSHCHA01
..
Align seg 1/1 to: HSSHCHA01 from: 1 to: 270
1 ArgGluAsnLeuArgThrAlaLeuArgTyr 10
|||||
222 CGAGAGAACCTGCGCACCGCTCCGCTAC 251

seq_name: gb_pr3:HSHLABS1

seq_documentation_block:
LOCUS HSHLABS1 270 bp DNA 30-OCT-1997
DEFINITION Human MHC class I antigen HLA-B GN00155-B*38022 gene, exon 2.
ACCESSION U90240
VERSION U90240.1 GI:1905853
KEYWORDS
SEGMENT
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
1 (bases 1 to 270)
AUTHORS Steiner,N.K., Ng,J., Bush,J., Hartzman,R.J., Johnson-Dow,L. and
Hurley,C.K.
TITLE HLA-B Alleles Associated with the B15 Serologically Defined
Antigens
JOURNAL Hum. Immunol. 56 (1997) In press
REFERENCE
2 (bases 1 to 270)
AUTHORS Steiner,N.K. and Hurley,C.K.
TITLE Direct Submission
JOURNAL Submitted (20-FEB-1997) Microbiology & Immunology, Georgetown
University, 3970 Reservoir Rd. NW, Washington, DC 20007, USA
FEATURES
Source
1..270
Location/Qualifiers
/organism="Homo sapiens"
/isolate="GN00155"
/db_xref="taxon:9606"
1..270
/gene="HLA-B"
/number=2
BASE COUNT 60 a 88 c 85 g 37 t
ORIGIN

exon
1..270
Quality: 50.00 Length: 10
Ratio: 5.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-25 x HSSHCHA01
..
Align seg 1/1 to: HSSHCHA01 from: 1 to: 270
1 ArgGluAsnLeuArgThrAlaLeuArgTyr 10
|||||
222 CGAGAGAACCTGCGCACCGCTCCGCTAC 251

seq_name: gb_pr3:HSHLABS1

seq_documentation_block:
LOCUS HSHLABS1 270 bp DNA 01-DEC-1997
DEFINITION Human HLA-B*15MD gene, exon 2.
ACCESSION U58315
VERSION U58315.1 GI:2654406
KEYWORDS
SEGMENT
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
1 (sites)
AUTHORS Gao,X. and Matheson,B.
TITLE A novel b15 variant found in oceanic populations

```



```

alignment_scores:
  Quality: 50.00      Length: 10
  Ratio: 5.000       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
  US-08-653-294-25 x HSHLABSH1 ..
  Align seg 1/1 to: HSHLABSH1 from: 1 to: 270
      1 ArgGluAsnLeuArgThrAlaLeuArgTyr 10
      |||||
      222 CGAGAGAACCTGCGACCGCTCGCTAC 251

seq_name: gb_pr4:AF017316

seq_documentation_block:
  LOCUS AF017316 270 bp DNA PRI 26-MAR-1999
  DEFINITION Homo sapiens MHC class I antigen (HLA-B) gene, partial cds.
  ACCESSION AF017316
  VERSION AF017316.1 GI:2394335
  KEYWORDS
  SOURCE human.
  ORGANISM Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
    Eutheria; Primates; Catarrhini; Hominidae; Homo.
  REFERENCE
    1 (bases 1 to 270)
    Cao, K., Burdett, L., Zhang, G. and Fernandez-Vina, M.
    Direct Submission
    Submitted (07-AUG-1997) Nat. Histocompatibility Lab, Am Red Cross,
    box 173, 22 S. Greene St., Baltimore, MD 21201, USA
  JOURNAL
  FEATURES
    source
      1..270
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /clone="27B"
      /chromosome="6"
      <1..>270
      /gene="HLA-B"
      /note="novel allele"
      <1..>270
      /gene="HLA-B"
      /codon_start=3
      /product="MHC class I antigen"
      /protein_id="AAB70282.2"
      /db_xref="GI:4521330"
      /translation="SISMRYFYTAMSRPGRPFRTVGVVDFTQVRFDSDATSPRM
      APRAPWIEQGPPEYWDRETOISKNTQTYRENLRALRYNQSEA"
  BASE COUNT 59 a 92 c 83 g 36 t
  ORIGIN

alignment_scores:
  Quality: 50.00      Length: 10
  Ratio: 5.000       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
  US-08-653-294-25 x AF017316 ..
  Align seg 1/1 to: AF017316 from: 1 to: 270
      1 ArgGluAsnLeuArgThrAlaLeuArgTyr 10
      |||||
      222 CGAGAGAACCTGCGACCGCTCGCTAC 251

seq_name: gb_pr4:HB38021G1

seq_documentation_block:
  LOCUS HB38021G1 270 bp DNA PRI 23-SEP-1999
  DEFINITION Homo sapiens HLA-B MHC class I antigen (HLA-B) gene, HLA-B-*38021
  allele, exon 2.
  ACCESSION AF181857

```

---

```

VERSION AF181857.1 GI:5919134
KEYWORDS
SEGMENT 1 of 2
SOURCE human.
ORGANISM Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
  Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 270)
  Pimtanonthai, N. and Hurley, C.K.
  Novel HLA-B allele
  Unpublished
  JOURNAL
REFERENCE
  2 (bases 1 to 270)
  Pimtanonthai, N. and Hurley, C.K.
  Direct Submission
  TITLE Submitted (30-AUG-1999) Microbiology & Immunology, Georgetown
  JOURNAL University, 3970 Reservoir Rd. N.W., Washington, D.C. 20007, USA
FEATURES
  source
    1..270
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone="49N-TA"
    <1..>270
    /gene="HLA-B"
    /number=2
  exon
  BASE COUNT 58 a 87 c 88 g 37 t
  ORIGIN

alignment_scores:
  Quality: 50.00      Length: 10
  Ratio: 5.000       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
  US-08-653-294-25 x HB38021G1 ..
  Align seg 1/1 to: HB38021G1 from: 1 to: 270
      1 ArgGluAsnLeuArgThrAlaLeuArgTyr 10
      |||||
      222 CGAGAGAACCTGCGACCGCTCGCTAC 251

seq_name: gb_pr4:HS44032S1

seq_documentation_block:
  LOCUS HS44032S1 270 bp DNA PRI 21-JUN-1999
  DEFINITION Human MHC class I antigen HLA-B (HLA-B*4410 allele) gene, exon 2.
  ACCESSION U63559
  VERSION U63559.1 GI:1488303
  KEYWORDS
  SEGMENT 1 of 2
  SOURCE human.
  ORGANISM Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
    Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 270)
  Hurley, C.K. and Hoyer, R.J.
  Human MHC Class I Antigen HLA-B*4410
  Unpublished
  JOURNAL
REFERENCE
  2 (bases 1 to 270)
  Hurley, C.K. and Hoyer, R.J.
  Direct Submission
  TITLE Submitted (11-JUL-1996) Dept. of Microbiology and Immunology,
  JOURNAL Georgetown University, 3970 Reservoir Rd., NW, Washington, DC
  20007, USA
FEATURES
  source
    1..270
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /chromosome="6"
    /cell_type="white blood cells from whole blood"
    1..270
    /gene="HLA-B"

```

BASE COUNT 60 a 89 c 85 g 36 t  
ORIGIN /number=2  
alignment\_scores: Quality: 50.00 Length: 10  
Ratio: 5.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-08-653-294-25 x HS44032S1 ..

Align seg 1/1 to: HS44032S1 from: 1 to: 270

1 ArgGlusnLeuargThraLeuArgTyr 10  
|||||  
222 CGAGAGAACTGCGCACCGCTCCGCTAC 251

seq\_name: gb\_pr4:AF035648

seq\_documentation\_block: 395 bp DNA PRI 05-JAN-1999  
LOCUS AF035648 Homo sapiens MHC class I antigen HLA-B gene (HLA-B\*4402 variant allele), partial cds.  
ACCESSION AF035648  
VERSION AF035648.1 GI:4104422  
KEYWORDS human.  
SOURCE  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 395)  
AUTHORS Holdsworth, R.  
TITLE Direct Submission  
JOURNAL Submitted (24-NOV-1997) Tissue Typing, Red Cross Blood Service, South Melbourne, Melbourne, Victoria 3205, Australia  
FEATURES  
source  
1..395  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="6"  
/gene="HLA-B"  
/product="MHC class I antigen"  
/protein\_id="AAD02035.1"  
/db\_xref="GI:4104423"  
/translation="FDSDATSPRKEPRAPWIEQGEPEYWDRETOISKNTQTYRENLR  
TALRYNQSAGSHIIQRMVGDVGPDRLLRGYDQDAYDKGYIALNEDLSWTAAD  
TAAQITQKWEARVAEQDRAYLEGEVE"  
BASE COUNT 89 a 120 c 141 g 45 t  
ORIGIN

alignment\_scores: Quality: 50.00 Length: 10  
Ratio: 5.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-08-653-294-25 x AF035648 ..

Align seg 1/1 to: AF035648 from: 1 to: 395

1 ArgGlusnLeuargThraLeuArgTyr 10  
|||||

120 CGAGAGAACTGCGCACCGCTCCGCTAC 149  
seq\_name: gb\_pr4:AF035649

seq\_documentation\_block: 395 bp DNA PRI 05-JAN-1999  
LOCUS AF035649 Homo sapiens MHC class I antigen HLA-B gene (HLA-B\*4403 variant allele), partial cds.  
ACCESSION AF035649  
VERSION AF035649.1 GI:4104424  
KEYWORDS human.  
SOURCE  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 395)  
AUTHORS Holdsworth, R.  
TITLE Direct Submission  
JOURNAL Submitted (24-NOV-1997) Tissue Typing, Red Cross Blood Service, South Melbourne, Melbourne, Victoria 3205, Australia  
FEATURES  
source  
1..395  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="6"  
/gene="HLA-B"  
/product="MHC class I antigen"  
/protein\_id="AAD02036.1"  
/db\_xref="GI:4104425"  
/translation="FDSDATSPRKEPRAPWIEQGEPEYWDPEITOISKNTQTYRENLR  
TALRYNQSAGSHIIQRMVGDVGPDRLLRGYDQDAYDKGYIALNEDLSWTAAD  
TAAQITQKWEARVAEQDRAYLEGEVE"  
BASE COUNT 87 a 122 c 139 g 47 t  
ORIGIN

alignment\_scores: Quality: 50.00 Length: 10  
Ratio: 5.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-08-653-294-25 x AF035649 ..

Align seg 1/1 to: AF035649 from: 1 to: 395

1 ArgGlusnLeuargThraLeuArgTyr 10  
|||||

120 CGAGAGAACTGCGCACCGCTCCGCTAC 149

seq\_name: gb\_pr2:HUMBW44HLA

seq\_documentation\_block: 404 bp DNA PRI 14-APR-1994  
LOCUS HUMBW44HLA Homo sapiens HLA-B\*44 variant.  
ACCESSION L31798  
VERSION L31798.1 GI:472361  
KEYWORDS human leukocyte antigen.  
SOURCE Homo sapiens blood DNA.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 404)  
AUTHORS Petersdorf, E.W.  
TITLE A Novel HLA-B\*44 Variant

JOURNAL Unpublished (1994)  
FEATURES  
Location/Qualifiers  
1. .404  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/cell\_type="leukocyte"  
/tissue\_type="blood"  
/map="chromosome 6p"  
misc\_feature  
1. .404  
variation  
203  
/note="This is a report of an HLA-B\*4402 variant  
(HLA-B\*4402v) that differs from the published sequence at  
base 203 (nucleotide position 75 of exon 3)."  
/phenotype="HLA-B\*44"  
/replace="t"  
BASE COUNT 90 a 123 c 142 g 49 t  
ORIGIN  
  
alignment\_scores:  
Quality: 50.00 Length: 10  
Ratio: 5.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000  
  
alignment\_block:  
US-08-653-294-25 x HUMBW44HLA ..  
Align seg 1/1 to: HUMBW44HLA from: 1 to: 404  
1 ArgGluAsnLeuArgThrAlaLeuArgTyr 10  
|||||  
80 CGAGAGAACCTGCGCACCGCGCTCCGCTAC 109

**THIS PAGE BLANK (USPTO)**

OM of: US-08-653-294-25 to: N\_Geneseq\_36:\* out\_format : pfs

Date: Feb 8, 2000 7:31 PM

About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp  
-O=/cnrl\_1/USPTO.spool/US08653294/runat\_04022000.160701\_15807/app\_query.fasta.2  
-DB=N\_Geneseq\_36 -QPM=fastap -SUFFIX=mg -GAPOP=12.000  
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000  
-CGAPEXT=0.500 -CGAPEXT=0.500 -XGAPEXT=10.000 -XGAPEXT=0.500  
-CGAPEXT=6.000 -CGAPEXT=7.000 -YGAPEXT=10.000 -YGAPEXT=0.500  
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR SCORE=pct  
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0  
-MAXLEN=1000000 -USER=US08653294 -NCPU=6 -ICPU=3 -NO\_XLPXY -WAIT  
-THREADS=1

Search information block:

Query: US-08-653-294-25

Query length: 10

Database: N\_Geneseq\_36:\*

Database sequences: 311585

Database length: 125096042

Search time (sec): 873.190000

score\_list:

Sequence	Strid	Orig	zScore	Escore	Len	Documentation
N_Geneseq_36:Q29167	+	44.00	160.43	0.2460	270	HLA-B*52 exon 2 alpha-1 domain
N_Geneseq_36:Q01834	+	44.00	148.10	1.20	1086	Sequence encoding HLA-B*51 anti
N_Geneseq_36:Q01822	+	44.00	148.10	1.20	1086	Sequence encoding HLA-B*52 anti
N_Geneseq_36:Q005693	+	44.00	148.08	1.20	1089	HLA-B*51 gene for production of
N_Geneseq_36:Q005701	+	44.00	148.08	1.20	1089	HLA-B*52 gene for production of
N_Geneseq_36:Q12114	+	44.00	148.08	1.20	1089	HLA-B*53 exon. HLA-B*53 gene,
N_Geneseq_36:N70935	+	40.00	133.52	7.76	1026	Sequence encoding the human h
N_Geneseq_36:N70935	+	40.00	131.76	35.08	3874	Sequence of genomic DNA encodi
N_Geneseq_36:N70935	+	40.00	117.10	63.73	6553	HLA B*27 consensus sequence. De
N_Geneseq_36:V21209_16	-	37.00	85.47	3.7e+03	64976	Continuation (17 of 17) of
N_Geneseq_36:V22706	-	35.00	108.13	201.36	2147	Mouse recombinase hSR2 gene
N_Geneseq_36:V35645	+	35.00	107.20	227.00	2386	Cladosporium oxysporum glucose
N_Geneseq_36:T70130	+	34.00	107.92	206.97	1437	Max-interacting protein coding
N_Geneseq_36:V73804	-	34.00	79.62	7.8e+03	35100	KSHV LTR DNA (nucleotides 70
N_Geneseq_36:V21209_00	+	34.00	69.50	2.8e+04	110000	Methanococcus jannaschii c
N_Geneseq_36:V19941	-	34.00	67.52	3.6e+04	137507	KSHV long unique coding regi
N_Geneseq_36:T47216	+	33.00	117.56	60.08	316	DNA sequence of IL33-C5a flank
N_Geneseq_36:T47217	+	33.00	115.95	73.87	379	DNA sequence of IL53-C5a flank
N_Geneseq_36:Q08098	+	33.00	109.14	177.00	818	PagC-pagD intergenic region. Sa
N_Geneseq_36:T70216	+	33.00	109.14	177.00	818	S. typhimurium pagC-pagD interg
N_Geneseq_36:Q090335	+	33.00	108.78	185.38	852	T. sergentii 33 kDa protein. New
N_Geneseq_36:Q055300	+	33.00	107.45	219.84	990	Theileria sergenti main piropl
N_Geneseq_36:X05748	+	33.00	105.82	270.95	1190	Nucleotide sequence of human H
N_Geneseq_36:V81737	+	33.00	104.62	315.85	1362	Spingomonas sp. strain AD109
N_Geneseq_36:X15191	+	33.00	99.90	578.40	2320	pag C. New anti-salmonella vac
N_Geneseq_36:Q26694	-	33.00	99.90	578.40	2320	pagC/AP fusion region. Detecti
N_Geneseq_36:Q27739	-	33.00	99.90	578.40	2320	pagC gene. Salmonella vaccine
N_Geneseq_36:Q085099	-	33.00	99.90	578.40	2320	S. typhimurium pagC gene. Att
N_Geneseq_36:T70212	+	33.00	96.16	935.33	3542	Human secreted protein gene 53
N_Geneseq_36:V59563	+	33.00	94.77	1.1e+03	4144	Nucleotide sequence of the Sph
N_Geneseq_36:X15190	+	33.00	92.58	1.5e+03	5305	Streptococcus pneumoniae genom
N_Geneseq_36:V22268	+	33.00	73.88	1.6e+04	43804	Chicken embryo lethal orphan
N_Geneseq_36:T86375	+	33.00	73.88	1.6e+04	43804	Complete genome sequence of a
N_Geneseq_36:X26690	+	33.00	65.73	4.5e+04	110000	Continuation (3 of 6) of V3
N_Geneseq_36:V30458_2	+	33.00	65.73	4.5e+04	110000	Continuation (3 of 6) of V3
N_Geneseq_36:V30459_2	+	33.00	65.73	4.5e+04	110000	Continuation (3 of 6) of V3
N_Geneseq_36:T60484	+	32.00	105.66	276.33	791	Rat calbindin cDNA sequence. En
N_Geneseq_36:T50947	+	32.00	102.04	439.86	1191	CAMP factor gene cfr. Nucleic
N_Geneseq_36:X14096	+	32.00	99.68	595.49	1555	H. pylori GHP1320 gene. New
N_Geneseq_36:X33945	+	32.00	97.20	817.81	2056	Human HCMV Inducible gene, SEQ
N_Geneseq_36:Q37687	+	32.00	95.33	1.0e+03	2540	Rat choline kinase gene. Rat d

N\_Geneseq\_36:X04639 + 32.00 93.90 1.2e+03 2985 ! Nucleic acid encoding an en  
N\_Geneseq\_36:X20565 - 32.00 93.23 1.4e+03 3219 ! Polynucleotide sequence fro  
N\_Geneseq\_36:V59714 - 32.00 91.79 1.6e+03 3791 ! Human secreted protein gene  
N\_Geneseq\_36:X20574 + 32.00 87.93 2.7e+03 5861 ! Polynucleotide sequence fro

seq\_name: N\_Geneseq\_36:Q29167

seq\_documentation\_block:

ID Q29167 standard; DNA; 270 BP.

AC Q29167;

DT 09-MAR-1993 (first entry)

DE HLA-B\*52 exon 2 alpha-1 domain.

KW Human leukocyte antigen; transgenic; germ cells; somatic cells;

KW expression; ss.

PN J04091731-A.

PF 25-MAR-1992.

PR 03-AUG-1990; 207329.

PR 03-AUG-1990; JP-207329.

PA (OLYU ) OLYMPUS OPTICAL CO.

DR WPI; 92-342893/42.

PT Transgenic non-human mammalian HLA-B\*52 gene - useful for

PT analysis of expression of gene structure, and prodn. of

PT mouse model of human disease

PS Disclosure; Fig 1; 8pp; Japanese.

CC The sequence shows the exon 2 alpha-1-domain of the human leukocyte

CC antigen-B\*52 gene. The complete gene may be introduced into non-

CC human mammals, pref. rat or mouse, or their ancestors at the primary

CC developmental biological step via transplantation into the zygote or

CC embryo to generate transgenic non-human mammals incorporating the

CC HLA-B\*52 gene in both their germ cells and somatic cells. Transgenic

CC non-human mammals contg. HLA-B\*52 are useful for the analysis of

CC expression of the gene, its structure, and prodn. of mouse models of

CC human disease. See also Q29166-72.

SQ Sequence 270 BP; 59 A; 88 C; 86 G; 37 T;

alignment\_scores:

Quality: 44.00 Length: 10

Ratio: 4.889 Gaps: 0

Percent Similarity: 90.000 Percent Identity: 90.000

alignment\_block:

US-08-653-294-25 x Q29167 ..

Align seg 1/1 to: Q29167 from: 1 to: 270

1 ArgGluAsnLeuArgThrAlaLeuArgTyr 10

|||||

222 CGAGAGAACTCGCGATCGCGCTCGCTAC 251

seq\_name: N\_Geneseq\_36:Q01834

seq\_documentation\_block:

ID Q01834 standard; DNA; 1086 BP.

AC Q01834;

DT 19-MAR-1991 (first entry)

DE Sequence encoding HLA-B\*51 antigen.

DE Probe: HLA class I DNA; immunogen; ss.

OS Homo sapiens.

PN EP354580-A.

PF 14-FEB-1990.

PR 10-AUG-1989.

PR 11-AUG-1988; JP-200758.

PA (OLYU) Olympus Optical Co., Ltd.

PI Kano K, Takiguchi;

DR WPI; 90-046289/07.

PT New DNA for class 1 human leukocyte antigens and derived probes and

PT transformed cells, useful for DNA typing, as immunogens etc.

PS Claim 1; Page 11; 23pp; English.

CC The HLA class I DNA can be used as a source of probes for use in DNA

CC typing. Transformed cells, which are useful as immunogens, can be

CC obtained by introducing these DNAs into eucaryotic cells.

SQ Sequence 1086 BP; 224 A; 334 C; 356 G; 172 T;

alignment\_scores:  
 Quality: 44.00 Length: 10  
 Ratio: 4.889 Gaps: 0  
 Percent Similarity: 90.000 Percent Identity: 90.000

## alignment\_block:

US-08-653-294-25 x Q01834 ..

Align seg 1/1 to: Q01834 from: 1 to: 1086

1 ArgGluAsnLeuArgThrAlaLeuArgTyr 10  
 |||||  
 294 CGAGAGAACCTGCGGATCGGCTCGCTAC 323

seq\_name: N\_Geneseq\_36:Q01822

## seq\_documentation\_block:

ID Q01822 standard; DNA; 1086 BP.

AC Q01822;  
 DT 19-MAY-1991 (first entry)  
 DE Sequence encoding HLA-Bw52 antigen.  
 KW Probe: HLA class I DNA; immunogen; ss.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT cds 1..1086  
 FT /\*tag= a

PN EP-354580-A.  
 PD 14-FEB-1990.  
 PF 10-AUG-1989.  
 PR 11-AUG-1988: JP-200758.  
 PA (OLYU) Olympus Optical Co., Ltd.  
 PI Kano K, Takiguchi;  
 DR WPI: 90-046289/07.  
 DR P-PSDB; R03142.  
 PT New DNA for class I human leucocyte antigens and  
 PT transformed cells, useful for DNA typing, as immunogens etc.  
 PS Claim 2: p11-12: 23pp; English.  
 CC The HLA class I DNA can be used as a source of probes for use in DNA  
 CC typing. Transformed cells, which are useful as immunogens, can be  
 CC obtained by introducing these DNAs into eucaryotic cells.  
 SQ Sequence 1086 BP; 223 A; 335 C; 358 G; 170 T;

alignment\_scores:  
 Quality: 44.00 Length: 10  
 Ratio: 4.889 Gaps: 0  
 Percent Similarity: 90.000 Percent Identity: 90.000

## alignment\_block:

US-08-653-294-25 x Q01822 ..

Align seg 1/1 to: Q01822 from: 1 to: 1086

1 ArgGluAsnLeuArgThrAlaLeuArgTyr 10  
 |||||  
 294 CGAGAGAACCTGCGGATCGGCTCGCTAC 323

seq\_name: N\_Geneseq\_36:Q05693

## seq\_documentation\_block:

ID Q05693 standard; DNA; 1089 BP.

AC Q05693;  
 DT 03-JAN-1991 (first entry)  
 DE HLA-B51 gene for production of monoclonal antibodies.  
 KW Allootype specific monoclonal anti-HLA antibodies; hybridomas;  
 KW transgenic animals; HLA-B51 gene; ss.  
 FH Key Location/Qualifiers  
 FT exon 1..73  
 FT /\*tag= a  
 FT /number=1  
 FT 74..343  
 FT exon 2

FT /\*tag= b  
 FT /number=2  
 FT /note="alpha 1-domain"  
 FT 344..619  
 FT /\*tag= c  
 FT /number=3  
 FT /note="alpha 2-domain"  
 FT 620..895  
 FT /\*tag= d  
 FT /number=4  
 FT /note="alpha 3-domain"  
 FT 896..1012  
 FT /\*tag= e  
 FT /number=5  
 FT 1013..1042  
 FT /\*tag= f  
 FT /number=6  
 FT 1043..1089  
 FT /\*tag= g  
 FT /number=7  
 PN EP-383183-A.  
 PD 22-AUG-1990.  
 PF 07-FEB-1990; 102424.  
 PR 08-FEB-1989; JP-Q29313.  
 PA (OLYU) OLYMPUS OPTICAL KK.  
 PI Takiguchi M;  
 DR WPI: 90-255479/34.  
 PT Allotype specific monoclonal anti- HLA antibodies prodn. - using  
 PT hybridomas derived from transgenic animals carrying HLA gene and  
 PT immunised with HLA antigen of different allotype  
 PS Disclosure; Fig 1 A-G; 20pp; English.  
 CC The human HLA-B51 gene was injected into fertilised mouse eggs and  
 CC then these introduced into the uterus of a pseudo pregnant mouse.  
 CC The young were tested to ensure incorporation of the gene into the  
 CC chromosome, and one of them mated 3 times with a normal male to  
 CC produce 16 young, seven of which carried the HLA-B51 gene.  
 CC The transgenic offspring were immunised with HLA antigen.  
 CC The spleen lymphocytes were fused with myeloma cells. Hybridomas  
 CC producing antibodies were selected.  
 CC See also Q05701.  
 SQ Sequence 1089 BP; 224 A; 335 C; 357 G; 173 T;

alignment\_scores:  
 Quality: 44.00 Length: 10  
 Ratio: 4.889 Gaps: 0  
 Percent Similarity: 90.000 Percent Identity: 90.000

## alignment\_block:

US-08-653-294-25 x Q05693 ..

Align seg 1/1 to: Q05693 from: 1 to: 1089

1 ArgGluAsnLeuArgThrAlaLeuArgTyr 10  
 |||||  
 295 CGAGAGAACCTGCGGATCGGCTCGCTAC 324

seq\_name: N\_Geneseq\_36:Q05701

## seq\_documentation\_block:

ID Q05701 standard; DNA; 1089 BP.

AC Q05701;  
 DT 03-JAN-1991 (first entry)  
 DE HLA-Bw52 gene for production of monoclonal antibodies.  
 KW Allootype specific monoclonal anti-HLA antibodies; hybridomas;  
 KW transgenic animals; HLA-Bw52 gene; ss.  
 FH Key Location/Qualifiers  
 FT exon 1..73  
 FT /\*tag= a  
 FT /number=1  
 FT 74..343  
 FT /\*tag= b  
 FT /number=2

```
FT exon /note="alpha 1-domain"
FT 344..619
FT /*tag= c
FT /number=3
FT /note="alpha 2-domain"
FT 620..895
FT /*tag= d
FT /number=4
FT /note="alpha 3-domain"
FT 896..1012
FT /*tag= e
FT /number=5
FT 1013..1042
FT /*tag= f
FT /number=6
FT 1043..1089
FT /*tag= g
FT /number=7
FT
FT EP-383183-A.
FT 22-AUG-1990.
FT 07-FEB-1990; 102424.
FT 08-FEB-1989; JP-029313.
FT (OLYU ) OLYMPUS OPTICAL KK.
FT Takiguchi M;
FT WPI; 90-255479/34.
FT Allotype specific monoclonal anti- HLA antibodies prodn. - using
FT hybridomas derived from transgenic animals carrying HLA gene and
FT immunised with HLA antigen of different allotype
FT Disclosure; Fig 1 A-G; 20pp; English.
FT The human HLA-Bw52 gene was introduced into mouse L cells and
FT then these cells used to immunise one of the transgenic mice
FT (See Q05693).
FT The spleen lymphocytes were fused with myeloma cells (P3x63-Ag8.653).
FT Hybridomas producing antibodies were selected.
FT Sequence 1089 BP; 223 A; 336 C; 359 G; 171 T;

alignment_scores:
  Quality: 44.00 Length: 10
  Ratio: 4.889 Gaps: 0
Percent Similarity: 90.000 Percent Identity: 90.000

alignment_block:
US-08-653-294-25 x Q05701 ..
Align seg 1/1 to: Q05701 from: 1 to: 1089

1 ArgGluAsnLeuArgThrAlaLeuArgTyr 10
|||||
295 CGAGAGACCTCGGATCGGCTCGCTAC 324

seq_name: N_Geneseq_36:Q12114

seq_documentation_block:
ID Q12114 standard; DNA; 1089 BP.
AC Q12114;
DT 29-AUG-1991 (first entry)
DE HLA-Bw53 exon.
KW Human leukocyte antigen; probe; major histocompatibility complex;
KW MHC; class I; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT cds 1..1089 /*tag= a
FT
FT J03112487-A.
FT 14-MAY-1991.
FT 22-SEP-1989; 247697.
FT 22-SEP-1989; JP-247697.
FT (OLYU ) OLYMPUS OPTICAL KK.
FT WPI; 91-182991/25.
FT P-PSDB; R12463.
FT HLA-Bw53 gene, DNA probe and transformant cells - used for
PT immunisation, identifying specificity of antiserum etc.
```

```
PS Claim 1; Page 1; lipp; Japanese.
CC Probes comprising part of the sequence can be used to identify
CC Class I genes. The DNA can be expressed for immunisation of
CC animals and prodn. of monoclonal antibodies specific for the
CC HLA-Bw53 antigen. See also J03112485 and J03112486. 174 T;
SQ Sequence 1089 BP; 222 A; 337 C; 356 G; 174 T;

alignment_scores:
  Quality: 44.00 Length: 10
  Ratio: 4.889 Gaps: 0
Percent Similarity: 90.000 Percent Identity: 90.000

alignment_block:
US-08-653-294-25 x Q12114 ..
Align seg 1/1 to: Q12114 from: 1 to: 1089

1 ArgGluAsnLeuArgThrAlaLeuArgTyr 10
|||||
295 CGAGAGACCTCGGATCGGCTCGCTAC 324

seq_name: N_Geneseq_36:N70935

seq_documentation_block:
ID N70935 standard; DNA; 1026 BP.
AC N70935;
DT 10-APR-1991 (first entry)
DE Sequence encoding the human histocompatibility antigen HLA B27.
KW Rheumatic disorder; genetic screening; diagnosis;
KW ankylosing spondylitis; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT cds 1..1026
FT
FT DE3542024-A.
FT 04-JUN-1987.
FT 28-NOV-1985; 542024.
FT 28-NOV-1985; DE-542024.
FT 21-DEC-1985; DE-545576.
FT (BEHW ) BEHRINGER AG.
FT Riethmuller G, Meo T, Weiss E, Szots H;
FT WPI; 87-157893/23.
FT P-PSDB; P70590.
FT DNA coding for antigen HLA B27 - and diagnostic reagents contg.
FT such DNA, antigen or antibody
PS Claim 2; Page 4; 5pp; German
CC The DNA may be used as a hybridisation probe for detecting the HLA
CC B27 gene, e.g. for assessing susceptibility to rheumatic disorders
CC such as ankylosis spondylitis, or may be used to transform cells
CC for prodn. of HLA B27. The HLA B27 may be used to detect HLA B27
CC antibody in human serum, or to produce mono- or polyclonal HLA B27
CC antibodies for use in immunoassay.
SQ Sequence 1026 BP; 213 A; 307 C; 344 G; 162 T;

alignment_scores:
  Quality: 40.00 Length: 10
  Ratio: 4.444 Gaps: 0
Percent Similarity: 90.000 Percent Identity: 80.000

alignment_block:
US-08-653-294-25 x N70935 ..
Align seg 1/1 to: N70935 from: 1 to: 1026

1 ArgGluAsnLeuArgThrAlaLeuArgTyr 10
|||||
223 CGAGAGACCTCGGACCTGCTCGCTAC 252

seq_name: N_Geneseq_36:N70225

seq_documentation_block:
ID N70225 standard; DNA; 3874 BP.
```

```

AC N70225;
DT 03-APR-1991 (first entry)
DE Sequence of genomic DNA encoding human histocompatibility antigen
DE HLA-B 27.
KW Ankylosing spondylitis; rheumatic disorder; diagnosis; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT intron 518..590
FT intron /*tag= a
FT intron 720..989
FT intron /*tag= b
FT intron 1090..1506
FT intron /*tag= c
FT intron 1932..2357
FT intron /*tag= d
FT intron 2450..2566
FT intron /*tag= e
FT intron 3009..3041
FT intron /*tag= f
FT intron 3148..3191
FT intron /*tag= g
PN EP-226069-A.
PD 24-JUN-1987.
PF 21-NOV-1986; 116139.
PR 01-JAN-1985; DE-542024.
PR 21-DEC-1985; DE-545576.
PA (BEHW ) BEHRINGERWERKE AG.
PI Szotz H, Weiss E, Dorner C, Lang M, Meo T, Riethmuller G;
DR WPI; 87-171469/25.
DR P-PSDB; P70155.
PT DNA coding for human histocompatibility antigen HLA-B 27 - useful
PT for diagnosis and antigen and antibody prodn.
PS Claim 1: p6; 13pp; German.
CC The DNA may be used to detect the HLA-B 27 gene (opt. mutated) in
CC human genetic material. The HLA-B 27 may be used to detect anti-HLA-
CC B 27 antibodies in human serum. The antibodies may be used to
CC determine HLA-B 27 levels in human serum, e.g. for diagnosis of
CC rheumatic disorders, esp. ankylosing spondylitis.
SQ Sequence 3874 BP; 751 A; 1094 C; 1171 G; 858 T;

alignment_scores:
Quality: 40.00 Length: 10
Ratio: 4.444 Gaps: 0
Percent Similarity: 90.000 Percent Identity: 80.000

alignment_block:
US-08-653-294-25 x N70225 ..
Align seg 1/1 to: N70225 from: 1 to: 3874

1 ArgGluAsnLeuArgThrAlaLeuArgTyr 10
|||||:|||||:|||||:|||||:|||||:
941 CGAGAGACCTCGCGACCTGCTCCGCTAC 970

seq_name: N_Geneseq_36:T61639

seq_documentation_block:
ID T61639 standard; DNA; 6553 BP.
AC T61639;
DT 05-JUN-1997 (first entry)
DE HLA B27 consensus sequence.
KW HLA B27; seronegative spondylarthropathy; ankylosing spondylitis;
KW Reiter's syndrome; arthritis; acute anterior uveitis; diagnosis;
KW ss; ds.
OS Homo sapiens.
FH Key Location/Qualifiers
FT mrna 3968..6653
FT intron /*tag= a
FT intron /*tag= b
FT intron /*note= "HLA-B27 3' flanking region, downstream of
FT intron 3' untranslated region"
FT intron 4112..4556
FT intron /*tag= b

```

```

FT /*note= "3' flanking region diagnostic for genetic
FT predisposition to SNSA"
FT 4270..4556
FT /*tag= b
FT /*note= "3' flanking region diagnostic for genetic
FT predisposition to SNSA"
FT misc_difference 4495
FT /*tag= d
FT /*note= "absence of cytosine at this site is
FT indicative of a predisposition to SNSA"
PN WO9709450-A1.
PD 13-MAR-1997.
PF 16-AUG-1996; U13256.
PR 01-SEP-1995; US-522942.
PA (CEDA-) CEDARS SINAI MEDICAL CENT.
PI Tyan DB;
DR WPI; 97-192924/17.
PT Detecting pre-disposition to seronegative spondylarthropathies -
PT from the absence of a C residue at a specific position in the
PT 3'-flanking region of the HLA B27 allele
PS Claim 1; Page 52-56; 68pp; English.
CC Genetic predisposition to seronegative spondylarthropathies (SNSA)
CC is detected by determining the absence of a cytosine nucleotide in
CC the 3' flanking region (see also T61647-48) of an HLA-B gene at a
CC position corresponding to nucleotide 4495 of the HLA-B27 consensus
CC sequence given in T61639. Probes and primers (see also T61640-46)
CC based on this region can be used in diagnostic assays to detect the
CC genetic predisposition to SNSA, and permit the distinction of B27+
CC individuals who are resistant to SNSA from B27+ normal individuals
CC who are susceptible (but as yet unaffected) to such diseases.
SQ Sequence 6553 BP; 1443 A; 1619 C; 2017 G; 1474 T;

alignment_scores:
Quality: 40.00 Length: 10
Ratio: 4.444 Gaps: 0
Percent Similarity: 90.000 Percent Identity: 80.000

alignment_block:
US-08-653-294-25 x T61639 ..
Align seg 1/1 to: T61639 from: 1 to: 6553

1 ArgGluAsnLeuArgThrAlaLeuArgTyr 10
|||||:|||||:|||||:|||||:|||||:
1102 CGAGAGACCTCGCGACCTGCTCCGCTAC 1131

seq_name: N_Geneseq_36:V21209_16

seq_documentation_block:
Continuation (17 of 17) of V21209 from base 1600001 (Methanococcus jannaschii circula
WP Sequence split into 17 fragments LOCUS V21209 Accession V21209
WP Fragment Name Begin End
WP V21209_00 1 110000
WP V21209_01 100001 210000
WP V21209_02 200001 310000
WP V21209_03 300001 410000
WP V21209_04 400001 510000
WP V21209_05 500001 610000
WP V21209_06 600001 710000
WP V21209_07 700001 810000
WP V21209_08 800001 910000
WP V21209_09 900001 1010000
WP V21209_10 1000001 1110000
WP V21209_11 1100001 1210000
WP V21209_12 1200001 1310000
WP V21209_13 1300001 1410000
WP V21209_14 1400001 1510000
WP V21209_15 1500001 1610000
WP V21209_16 1600001 1664976

alignment_scores:

```



Quality: 37.00 Length: 10  
Ratio: 3.700 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 70.000

## alignment\_block:

US-08-653-294-25 x V21209\_16/rev ..

Align seg 1/1 to reverse of: V21209\_16 from: 1 to: 64976

1 ArgGluAsnLeuArgThrAlaLeuArgTyr 10  
49695 AGAGAAATTTAAAGACAGTCTAAGATAT 49666

seq\_name: N\_Geneseq\_36:V22706

## seq\_documentation\_block:

ID V22706 standard; DNA; 2147 BP.

AC V22706;  
DE 28-AUG-1998 (first entry)  
KW Mouse recombinase hREC2 gene promoter.  
KW Recombinase; REC2; hREC2; homologous recombination; irradiation;  
KW apoptosis; cancer; human; promoter; ss.  
OS Mus musculus.

FH Key Location/Qualifiers

FT misc\_feature 4..13  
FT misc\_feature 138..145  
FT misc\_feature 144..155  
FT misc\_feature 304..312  
FT misc\_feature 363..370  
FT misc\_feature 363..370  
FT misc\_feature 436..444  
FT misc\_feature 449..459  
FT misc\_feature 500..506  
FT misc\_feature 749..805  
FT misc\_feature 822..833  
FT misc\_feature 981..988  
FT misc\_feature 1119..1127  
FT misc\_feature 1119..1127  
FT misc\_feature 1219..1225  
FT misc\_feature 1262..1272  
FT misc\_feature 1434..1445

/\*tag= a  
/label= RNR3  
/\*tag= b  
/label= RAD23  
/\*tag= c  
/label= RAD6  
/\*tag= d  
/label= RHP51  
/\*tag= e  
/label= RNR3  
/\*tag= f  
/label= RAD51-UAS  
/\*tag= g  
/label= RAD7  
/\*tag= h  
/label= RAD54-UAS  
/\*tag= i  
/label= RAD-1  
/\*tag= j  
/label= PHR1-1  
/\*tag= k  
/label= PHR1-UAS  
/\*tag= l  
/label= RHP51-UAS  
/\*tag= m  
/label= RNR3-1  
/\*tag= n  
/label= RAD16-1  
/\*tag= o  
/label= RAD16  
/\*tag= p  
/label= DRE2

/\*tag= q  
/label= PHR1-UAS  
1533..1543  
/\*tag= r  
/label= RHP51-UAS  
1561..1567  
/\*tag= s  
/label= RHP51-UAS  
1584..1599  
/\*tag= t  
/label= UAS  
1699..1707  
/\*tag= u  
/label= RAD2-UAS  
1884..1892  
/\*tag= v  
/label= RHP51-2  
1895..1901  
/\*tag= w  
/label= PHR1  
1951..1956  
/\*tag= x  
/label= RHP51-UAS  
1954..1963  
/\*tag= y  
/label= RHP51-UAS  
2035..2048  
/\*tag= z  
/label= DRS

WO9811214-Al.  
19-MAR-1998.  
PD 11-SEP-1997; IB1217.  
PR 11-SEP-1996; US-025929.  
PA (CORR ) CORNELL RES FOUND INC.  
PA (UIJE-) UNIV JEFFERSON THOMAS.  
PI Holloman WK, Kmiec EB, Rice MC, Shu Z, Smith ST;  
DR WPI; 98-207378/18.  
PT - useful for the sensitisation of cells to the apoptotic effects of irradiation  
PS Disclosure; Fig 7D-F; 88pp; English.  
CC This polynucleotide comprises the promoter region of muREC2 (see V22704), the murine homologue of the REC2 gene of *Ustilago maydis* that codes for an ATP dependent recombinase (see W56265). The muREC2 promoter was sequence by digestion of clone lambda 5D2a with XbaI. It contains a number of motifs that show homology to radiation-sensitive cis-acting control sequences from yeast genes. The human REC2 gene (hREC2) promoter (see V22705) is also disclosed. The REC2 promoter, in combination with a strong enhancer, e.g. a SV40 enhancer, was found to be a strong promoter following irradiation of cells. A radiation inducible promoter can be used to sensitise a cell to radiation treatment by operably linking the promoter to a gene whose expression converts a prodrug to a drug, e.g. the herpes simplex virus thymidine kinase gene. SQ Sequence 2147 BP; 570 A; 349 C; 453 G; 775 T;

## alignment\_scores:

Quality: 35.00 Length: 10  
Ratio: 4.375 Gaps: 0  
Percent Similarity: 80.000 Percent Identity: 70.000

## alignment\_block:

US-08-653-294-25 x V22706/rev ..

Align seg 1/1 to reverse of: V22706 from: 1 to: 2147

1 ArgGluAsnLeuArgThrAlaLeuArgTyr 10  
432 AGAGAAATTTACACATTTATACGCTAC 403

seq\_name: N\_Geneseq\_36:V35645

```

seq_documentation_block:
ID V35645 standard; DNA; 2386 BP.
AC V35645;
DE 08-SEP-1998 (first entry)
DE Cladosporium oxysporum glucose oxidase encoding DNA.
KW Glucose oxidase; Cladosporium oxysporum; enzyme; bread improver;
KW antimicrobial agent; toothpaste; detergent; stain removal;
KW dough additive; hydrogen peroxide generator; ss.
OS Cladosporium oxysporum
Key Location/Qualifiers
FT 5'UTR 1..350
FT CDS 351..2192
FT FT /*tag= a
FT FT /*tag= b
FT FT /product= "glucose oxidase"
FT FT 351..419
FT FT /*tag= c
FT FT /note= "predicted signal sequence"
FT FT mat_peptide 420..2189
FT FT /*tag= d
FT FT 3'UTR 2190..2386
FT FT /*tag= e
PN W09820136-A1.
PD 14-MAY-1998.
PF 03-NOV-1997; U20174.
PF 07-NOV-1996; US-746257.
PA (NOVO ) NOVO NORDISK BIOTECH INC.
PA (NOVO ) NOVO NORDISK AS.
PI Berka RM, Cherry JR, Halkier T;
DR WPI; 98-286952/25.
DR P-PSDB; W60593.
PT New nucleic acid encoding glucose oxidase active at acidic pH, from
PT Cladosporidium - and related vectors and host cells, producing
PT enzyme useful as bread improver, antimicrobial additive for
PT toothpaste, etc. and hydrogen peroxide generator in detergents
PS Claim 4; Fig 2A-B; 83pp; English.
CC This DNA encodes a Cladosporium oxysporum glucose oxidase. Host cells
CC containing a construct comprising the glucose oxidase encoding nucleic
CC acid sequence with regulatory sequences are used to produce recombinant
CC glucose oxidase. Preferred fragments of the nucleic acid are present in
CC pGEX4A and pGEX6A, contained in E. coli deposited as NREL B-21628 and
CC B-21629. The host cells are particularly Fusarium. The recombinant
CC glucose oxidase is useful as dough additive to improve the gluten
CC quality. It can be used as additive for toothpaste (particularly used
CC with a thiocyanate and lactoperoxidase to generate antimicrobial
CC oxythiocyanate anion), mouthwash, denture cleaners, soaps, hair and body
CC care products. It can be added to cleaning solutions for contact lenses
CC as bleach or as an antibacterial agent. It is also useful as a hydrogen
CC peroxide generator in laundry and dishwashing detergents, particularly
CC for stain removal.
SQ Sequence 2386 BP; 583 A; 707 C; 581 G; 515 T;

alignment_scores:
Quality: 35.00 Length: 9
Ratio: 3.889 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 77.778

alignment_block:
US-08-653-294-25 x V35645 ..
Align seg 1/1 to: V35645 from: 1 to: 2386

1 ArgGluAsnLeuArgThrAlaLeuArg 9
|||||:|||||:|||||:|||||:
605 COTGAGGACTTACGGCAAGCCTCGA 631

seq_name: N_Geneseq_36:T70130

seq_documentation_block:
ID T70130 standard; cDNA to mRNA; 1437 BP.
AC T70130;
DE 07-FEB-1998 (first entry)

```

```

DE Max-interacting protein coding sequence (clone 18).
KW murine; mSinA; mammalian homologue; Saccharomyces cerevisiae; repressor;
KW Sin3; Mad; Max; mSin; Mad complex; mSin; Mad:Max complex; Myc; promoter;
KW basic helix-loop-helix zipper protein; compete; DNA-binding;
KW Myc:Max complex; activate; transcription; gene regulation; ss.
OS Mus musculus.
PN US5624818-A.
PD 29-APR-1997.
PD 01-JUN-1994; 252966.
PD 01-JUN-1994; US-252966.
PR 19-SEP-1991; US-756195.
PR 23-JUN-1992; US-903710.
PR 01-APR-1994; US-222638.
PA (HUTC-) HUTCHINSON CANCER RES CENT FRED.
PA Ayer DE, Eisenman RN;
PI WPI; 97-258216/23.
PT mSin nucleic acids encoding recombinant polypeptide(s) that
PT associate with Mad polypeptide - are possible homologues of S.
PT cerevisiae general repressor protein Sin3
PS Example 13; Fig 25; 111pp; English.
CC This cDNA sequence, designated clone 18, encodes a novel murine basic
CC helix-loop-helix zipper (bHLHZip) protein which interacts with the
CC transcriptional functions of Myc family proteins as well as for the Mad
CC protein. Max is a stable, ubiquitously expressed protein which in general
CC does not appear to be regulated during mitogenesis, the cell cycle, or
CC differentiation. Other cDNA sequences encoding murine proteins (msin)
CC which may be mammalian homologues of the Saccharomyces cerevisiae general
CC repressor protein Sin3 are claimed. The mSin proteins associate with Mad
CC and Max to form a mSin:Max complex which binds to a CACGTG sequence
CC in promoters. Mad:Max complexes repress, while Myc:Max complexes
CC activate, transcription from promoters containing proximal CACGTG binding
CC sites for these proteins. Expression of Mad is closely linked to
CC differentiation in at least two distinct cell lineages. The switch from
CC Myc:Max to Mad:Max complexes may reflect the repression of transcription
CC of Myc regulated genes by Mad. The DNA, vectors and host cells of the
CC invention are useful for the recombinant production of mSin proteins
CC useful in elucidation of Mad repressor functions.
SQ Sequence 1437 BP; 302 A; 532 C; 392 G; 211 T;

alignment_scores:
Quality: 34.00 Length: 7
Ratio: 4.857 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-25 x T70130 ..
Align seg 1/1 to: T70130 from: 1 to: 1437

4 LeuArgThrAlaLeuArgTyr 10
|||||:|||||:|||||:|||||:
889 CTGCGGAGCGGCTGCGGTAC 909

seq_name: N_Geneseq_36:T73804

seq_documentation_block:
ID V73804 standard; DNA; 35100 BP.
AC V73804;
DE 25-FEB-1999 (first entry)
DE KSHV LUR DNA (nucleotides 70,201-105,300).
KW Kaposi's sarcoma; acquired immune deficiency syndrome; AIDS; DHFR; Bcl-2;
KW dihydrofolate reductase; LUR; long unique region; vaccine; prophylaxis;
KW diagnosis; treatment; HHV8; transactivator; LCTP; dUTPase; Isp-II;
KW DNA replication protein I; immediate early protein II; vRFL1; ICSBP;
KW phosphoprotein; DNA replication protein II; ribonucleotide reductase;
OS assembly/DNA maturation protein; tegument protein; ds.
OS Kaposi's sarcoma-associated herpesvirus.
PN US5849564-A.
PD 15-DEC-1998.
PF 29-NOV-1996; 770379.
PF 29-NOV-1996; US-770379.

```

PA (UYCO ) UNIV COLUMBIA NEW YORK.  
 PI Bohenzky RA, Chang Y, Edelman IS, Moore PS, Russo JJ;  
 DR WPI: 99-069741/06.  
 PT Kaposi's sarcoma-associated herpes virus nucleic acid - encodes  
 PT di:hydrofolate reductase and is useful for treatment, prophylaxis  
 PT or diagnosis of Kaposi's sarcoma  
 PS Disclosure: Column 125-156; 109pp; English.  
 CC This sequence is a fragment of the Kaposi's sarcoma-associated  
 CC herpesvirus (KSHV) LUR (long unique region). This fragment contains  
 CC coding regions for ORF48, ORF49, ORF50 which encodes a transactivator  
 CC (LCTP), A8, ORF52, ORF53, ORF54 which encodes dUTPase, ORF55, ORF56  
 CC which encodes DNA replication protein I, ORF57 which encodes immediate  
 CC early protein II (IEP-II), K9 which encodes virf1 (ICSBP), K10, K11,  
 CC ORF58 which encodes a phosphoprotein, ORF59 which encodes DNA replication  
 CC protein II, ORF60 which encodes a small ribonucleotide reductase, ORF61  
 CC which encodes a large ribonucleotide reductase, ORF62 which encodes an  
 CC assembly/DNA maturation protein, ORF63 which encodes tegument protein  
 CC II, ORF64 which encodes tegument protein III. KSHV is a new human  
 CC Herpesvirus (HHV8) believed to cause Kaposi's sarcoma (KS) which is the  
 CC most common form of neoplasm occurring in persons with acquired immune  
 CC deficiency syndrome (AIDS). The DHFR protein is useful for vaccination,  
 CC prophylaxis, diagnosis and treatment of a subject with Kaposi's sarcoma  
 CC and for detecting expression of a DNA virus associated with Kaposi's  
 CC sarcoma in a cell.  
 SQ Sequence 35100 BP; 8632 A; 10010 C; 8650 G; 7808 T;

alignment\_scores:  
 Quality: 34.00 Length: 9  
 Ratio: 4.250 Gaps: 0  
 Percent Similarity: 88.889 Percent Identity: 77.778

alignment\_block:  
 US-08-653-294-25 x V73804/rev ..  
 Align seg 1/1 to reverse of: V73804 from: 1 to: 35100

1 ArgGluAsnLeuArgThrAlaLeuArg 9  
 3575 CGGGAGATCTCCGGTGTCTCTGCGA 3549

seq\_name: N\_Geneseq\_36:V21209\_00

seq\_documentation\_block:  
 WP Sequence split into 17 fragments LOCUS V21209 Accession V21209  
 WP Fragment Name Begin End  
 WP V21209\_00 1 110000  
 WP V21209\_01 100001 210000  
 WP V21209\_02 200001 310000  
 WP V21209\_03 300001 410000  
 WP V21209\_04 400001 510000  
 WP V21209\_05 500001 610000  
 WP V21209\_06 600001 710000  
 WP V21209\_07 700001 810000  
 WP V21209\_08 800001 910000  
 WP V21209\_09 900001 1010000  
 WP V21209\_10 1000001 1110000  
 WP V21209\_11 1100001 1210000  
 WP V21209\_12 1200001 1310000  
 WP V21209\_13 1300001 1410000  
 WP V21209\_14 1400001 1510000  
 WP V21209\_15 1500001 1610000  
 WP V21209\_16 1600001 1664976  
 ID V21209 standard: DNA; 1664976 BP.  
 AC V21209;  
 DT 10-NOV-1998 (first entry)  
 DE Methanococcus jannaschii circular chromosome.  
 DE Methanococcus jannaschii; methanogenic archaeon; circular chromosome;  
 KW genome; autotrophic; extrachromosomal element; identification; ds.  
 OS Methanococcus jannaschii.  
 PN W09807830-A2.  
 PD 26-FEB-1998.  
 PF 22-AUG-1997; U14900.

PR 22-AUG-1996; US-024428.  
 PA (GENO-) INST GENOMIC RES.  
 PA (UNII ) UNIV ILLINOIS FOUND.  
 PA (UYJO ) UNIV JOHNS HOPKINS SCHOOL MEDICINE.  
 PI Sult CJ, Smith HO, Venter JC, White OR, Woese CR;  
 DR WPI: 98-169145/15.  
 PT Complete genome sequence of methano-genic archaeon, Methanococcus  
 PT jannaschii - useful in identification of M. jannaschii genome  
 PT fragment  
 PS Claim 13: Page 152-585; 614pp; English.  
 CC The present sequence represents the complete 1.66-megabase pair genome  
 CC sequence of the Methanococcus jannaschii circular chromosome. The  
 CC present invention describes M. jannaschii open reading frames from the  
 CC genome sequence. The invention also describes a computer based system  
 CC for identifying fragments of the M. jannaschii genome that are  
 CC homologous to target nucleotide sequences, comprising: (a) data storage  
 CC means comprising the nucleotide sequence of the 1664976, 58407 or 16550  
 CC bp sequence (see V21209, V21210 and V21211), or a nucleotide sequence at  
 CC least 99.9% identical to it; (b) search means for comparing a target  
 CC sequence to the nucleotide sequence of the data storage means to  
 CC identify a homologous sequence, and (c) retrieval means for obtaining  
 CC the homologous sequence. The method, which is based on whole genome  
 CC random sequencing of an autotrophic archaeon M. jannaschii, the genome  
 CC of which consists of 3 physically distinct elements, a large circular  
 CC chromosome (the 1664976 bp sequence given in V21209), a large circular  
 CC extra-chromosomal element (the 58407 bp sequence given in V21210), and a  
 CC small circular extra-chromosomal element (the 16550 bp sequence given in  
 CC V21211), can be used in the identification of M. jannaschii genome  
 CC fragment.  
 SQ Sequence 1664976 BP; 568133 A; 264649 C; 258701 G; 573392 T;

alignment\_scores:  
 Quality: 34.00 Length: 9  
 Ratio: 4.250 Gaps: 0  
 Percent Similarity: 88.889 Percent Identity: 66.667

alignment\_block:  
 US-08-653-294-25 x V21209\_00 ..

Align seg 1/1 to: V21209\_00 from: 1 to: 110000

2 GluAsnLeuArgThrAlaLeuArgTyr 10  
 |||||:::|||||:::|||||  
 94200 GAGAAATTTAAACCTCTACTAAATAT 94226

**THIS PAGE BLANK (USPTO)**

OM of: US-08-653-294-25 to: EST:\* out\_format: pfs  
 Date: Feb 8, 2000 6:23 AM  
 About: Results were produced by the GenCore software, version 4.5,  
 Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:  
 -MODEL=frame p2n.model -DEV=xlp  
 -O=/cnl1/USPTO.spool/US08653294/runat\_04022000\_160700\_15770/app\_query.fasta.2  
 -DB=EST -QFMT=fastap -SUFFIX=rst -GAPOP=12.000 -GAPEXT=4.000  
 -MINMATCH=0.100 -LOOPCL=0.000 -LOPEXT=0.000 -QGAPOP=4.500  
 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000  
 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000  
 -DELEXT=7.000 -START=1 -MATRIX=bscsum62 -TRANS=human40.cdi  
 -LIST=45 -DOCALLIGN=200 -MTRX=score-pct -ALIGN=15 -MODE=LOCAL  
 -OUTFMT=pfs -NORM=ext -MINLEN=0 -MAXLEN=1000000 -USER=US08653294  
 -NCPU=6 -ICPU=3 -NO\_XLPXY -WAIT -THREADS=1

# Search information block:

Query: US-08-653-294-25  
 Query length: 10  
 Database: EST.\*  
 Database sequences: 4538634  
 Database length: 1887831982  
 Search time (sec): 7600.090000

score_list:	Sequence	Strd Orig	ZScore	EScore Len	Documentation
gb_est23:AI124815	+	50.00	185.58	0.1431	AI124815 am5606.x1 Johnston fr
gb_est8:CO3945	+	44.00	167.16	1.52	AI3945 CO3945 Human heart CDNA
gb_est10:AA151891	+	44.00	166.30	1.70	AA151891 z001f06.r1 Stratagene
gb_est11:AA263158	+	44.00	165.34	1.92	AA263158 PMY0534 KGI-a Lambda Z
gb_est6:D82221	+	44.00	162.77	2.67	D82221 HUMHBC4626 Human pancrea
gb_est16:AI159260	-	44.00	158.20	4.79	AI159260 gy27b07.x1 NCI_CGAP_Br
gb_est31:AI169686	-	44.00	156.45	6.00	AI169686 wC74h11.x1 NCI_CGAP_Pa
gb_est27:AI466429	+	43.00	159.04	4.30	AI466429 vx35b04.y1 Stratagene
gb_est20:AA881004	+	43.00	154.92	7.30	AA881004 vx35b04.r1 Stratagene
gb_gss7:AO894863	-	41.00	148.17	17.35	AO894863 HS_3193_A2.D11.T7C CIT
gb_est18:AI036690	+	40.00	144.14	8.06	AI036690 DKFp564D2463.r1 564 C
gb_est10:AA147151	+	40.00	142.95	33.87	AA147151 z032d06.r1 Stratagene
gb_est37:AI197388	-	39.00	142.76	34.69	AI197388 sd11c05.y1 Gm-cl020 GI
gb_est18:AA722000	-	39.00	141.38	41.45	AA722000 zhl7e08.s1 Soares.pine
gb_est5:AO64631	-	39.00	141.30	41.88	AO64631 zf72c03.s1 Soares.pine
gb_est9:AO64793	+	39.00	141.02	43.41	AO64793 zf72c03.r1 Soares.pine
gb_est6:D82189	+	38.00	138.12	62.90	D82189 HUMHBC4524 Human pancrea
gb_est44:AW212923	+	38.00	136.17	80.85	AW212923 uc069c04.x1 NCI_CGAP_Ma
gb_est44:AW215398	+	38.00	134.89	95.24	AW215398 uc03c01.x1 NCI_CGAP_Lu
gb_est32:AI007865	-	37.00	136.41	78.38	AI007865 EST202316 Normalized t
gb_est14:AA445358	-	37.00	135.30	90.40	AA445358 vf64d07.r1 Barstede ME
gb_est23:AI089530	-	37.00	134.98	94.17	AI089530 z222a07.x1 Soares.tota
gb_est22:AI066062	-	37.00	133.27	117.23	AI066062 BSMHFS209D/SK Brugia
gb_est2:RI3904	+	37.00	133.25	117.53	RI3904 yf62c03.r1 Soares infant
gb_est16:AA570443	-	37.00	132.94	122.36	AA570443 qa64a02.x1 Soares.feta
gb_gss11:AO275243	-	37.00	132.78	124.78	AO275243 RPI-1-1319T7 RPCI-1 H
gb_est21:AA910336	+	37.00	132.75	125.38	AA910336 kc83d08.s1 NCI_CGAP_Ki
gb_est26:AA817822	+	37.00	131.85	140.69	AA817822 UI-rA0-aed-07-0-UI-8
gb_gss3:B60318	-	37.00	131.76	142.23	B60318 CIT-HSP-2378B15.TF CIT
gb_gss8:AO20193	-	37.00	131.36	145.95	AO20193 CIT-HSP-2306N22.TR CIT
gb_gss12:AO366195	-	37.00	130.97	157.49	AO366195 nbx0065M10f CUGI Rice
gb_est42:AV324628	+	36.00	134.71	97.45	AV324628 AV324628 RIKEN full-l
gb_gss9:AO111757	-	36.00	131.91	139.58	AO111757 CIT-HSP-2378B15.TF CIT
gb_gss14:AO511513	+	36.00	130.04	177.30	AO511513 HS_3074_B1.C12.SPES RH
gb_gss15:AO614324	+	36.00	128.17	225.39	AO614324 HS_5126_B1.D10.T7A RQC
gb_gss9:AO164696	+	36.00	128.12	226.92	AO164696 HS_3006_B2.D07.T7 CIT
gb_gss10:AO218520	-	36.00	128.10	227.43	AO218520 HS_3247_B1.E10.T7 CIT
gb_gss13:AO459514	-	36.00	127.96	231.52	AO459514 HS_5085_A1.B10.T7A RQC
gb_est39:AW132450	-	36.00	127.54	244.36	AW132450 se03e07.y1 Gm-cl013 GI
gb_gss14:AO570343	-	36.00	127.44	247.46	AO570343 HS_5352_B1.G08.T7A RQC
gb_gss11:AO328708	+	36.00	127.17	256.26	AO328708 nbx00043P15f CUGI Rice

gb\_gss11:AO322213 + 36.00 126.94 264.07 594 ! AO322213 RPCI11-96G8.TJ RPCI  
 gb\_gss3:B90327 - 36.00 125.88 302.55 667 ! B90327 CIT-HSP-2173G20.TF CI  
 gb\_est19:AA541035 + 36.00 125.84 304.15 670 ! AA541035 LD20818.5prime LD D

seq\_name: gb\_est23:AI124815

## seq\_documentation\_block:

LOCUS AI124815 414 bp mRNA EST 11-SEP-1998  
 DEFINITION am5606.x1 Johnston frontal cortex Homo sapiens cDNA clone  
 IMAGE:1539596 3' similar to gb:M24038 cds1 HLA CLASS I  
 HISTOCOMPATIBILITY ANTIGEN, BW-44(B-12) B\*4402 (HUMAN); contains  
 MER22.t3 TARI repetitive element ; mRNA sequence.

ACCESSION AI124815

VERSION AI124815

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 414)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,  
 Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M.,  
 Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F.,  
 Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.

TITLE WashU-NCI human EST Project

JOURNAL Unpublished (1997)

COMMENT On Jan 17, 1998 this sequence version replaced gi:1899887.

Contact: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@wustl.edu

This clone is available royalty-free through LLNL : contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Seq primer: -40ml3 fwd. Et from Amersham.

## FEATURES

Location/Qualifiers

1..414

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:1539586"

/clone\_lib="Johnston frontal cortex"

/sex="male"

/tissue\_type="pooled frontal lobe"

/dev\_stage="adult"

/lab\_host="SOLR (kanamycin resistant)"

/note="Organ: brain; Vector: Bluescript SK-; Site\_1:

ECORI; Stanley Neurophatology Consortium

(www.stanleylab.org) brains S-58, S-65, S-67, S-78.

Random + oligo-dt primed into EcorI site of ZAP II Vector.

Mass excised. Avg insert length 1.9kb. Custom library

provided by Dr. Nancy Johnston [(410) 614-3918,

nj@welchlink.welch.jhu.edu].

BASE COUNT 80 a 140 c 136 g 58 t

## ORIGIN

alignment\_scores:

Quality: 50.00

Ratio: 5.000

Percent Similarity: 100.000

Percent Identity: 100.000

Length: 10

Gaps: 0

US-08-653-294-25 x AI124815

Align seg 1/1 to: AI124815 from: 1 to: 414

1 ArgGluAsnLeuArgThrAlaLeuArgTyr 10

|||||

284 CGAGAGACCTGCGACCGCGCTCCGCTAC 313

seq\_name: gb\_est8:CO3945

```

seq_documentation_block: 232 bp mRNA EST 30-JUL-1996
LOCUS C03945 Human heart cDNA (Ynakamura) Homo sapiens CDNA clone
DEFINITION 3NHC2454, mRNA sequence.
ACCESSION C03945
VERSION C03945.1 GI:1467196
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 232)
AUTHORS Tanaka,T., Ogiwara,A., Uchiyama,I., Takagi,T., Yazaki,Y. and Nakamura,Y.
TITLE Construction of a normalized directionally cloned cDNA library from adult heart and analysis of 3040 clones by partial sequencing
JOURNAL Genomics 35 (1), 231-235 (1996)
MEDLINE
COMMENT On Oct 24, 1995 this sequence version replaced gi:1040105.
Contact: Yusuke Nakamura
Institute of Medical Science
University of Tokyo
4-6-1, Shirokanedai, Minato-ku, Tokyo 108, Japan
Tel: 81-3-5449-5372
Fax: 81-3-5449-5433
Email: yusuke@ims.u-tokyo.ac.jp.
FEATURES
Location/Qualifiers
source 1..232
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="3NHC2454"
/clone_lib="Human heart cDNA (Ynakamura)"
/dev_stage="adult"
/note="organ: heart; normalized directionally cloned cDNA from adult heart"
BASE COUNT 55 a 77 c 68 g 32 t
ORIGIN

alignment_scores:
Quality: 44.00 Length: 10
Ratio: 4.889 Gaps: 0
Percent Similarity: 90.000 Percent Identity: 90.000

alignment_block:
US-08-653-294-25 x C03945
Align seg 1/1 to: C03945 from: 1 to: 232
1 ArgGluAsnLeuArgThrAlaLeuArgTyr 10
|||||
40 CGAGAGAACCTGCGGATCGGCTCGCTAC 69

seq_name: gb_est10:AA151891

seq_documentation_block: 255 bp mRNA EST 10-DEC-1996
LOCUS AA151891
DEFINITION z001f06.r1 Stratagene colon (#937204) Homo sapiens CDNA clone IMAGE:566435 5' similar to gb:ML5497_cds1_HLA_CLASS_I HISTOCOMPATIBILITY ANTIGEN, A-24(A-9) A*2401 (HUMAN), mRNA sequence.
ACCESSION AA151891
VERSION AA151891.1 GI:1720754
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 255)
AUTHORS Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B., Chisoe,S., Dietrich,N., Dubuque,T., Favellio,A., Gish,W., Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L.,

```

```

Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Mieg,J.,
Trevaskis,E., Underwood,K., Wohldmann,P., Waterston,R., Wilson,R.
and Marra,M.
Generation and analysis of 280,000 human expressed sequence tags
Genome Res. 6 (9), 807-828 (1996)
97044478
On May 8, 1995 this sequence version replaced gi:800234.
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Trace considered overall poor quality
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 1.
Location/Qualifiers
source 1..255
/organism="Homo sapiens"
/db_xref="GDB:450888"
/db_xref="taxon:9606"
/clone="IMAGE:566435"
/clone_lib="Stratagene colon (#937204)"
/lab_host="SOLR cells (kanamycin resistant)"
/note="Organ: colon; Vector: pBluescript SK-; Site: 1: EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer: Oligo dT. T-84 colonic epithelial cell line. Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGACGAG 3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTT 3'"
BASE COUNT 57 a 70 c 75 g 44 t 9 others
ORIGIN

alignment_scores:
Quality: 44.00 Length: 10
Ratio: 4.889 Gaps: 0
Percent Similarity: 90.000 Percent Identity: 90.000

alignment_block:
US-08-653-294-25 x AA151891
Align seg 1/1 to: AA151891 from: 1 to: 255
1 ArgGluAsnLeuArgThrAlaLeuArgTyr 10
|||||
77 CGAGAGAACCTGCGGATCGCTCGCTAC 106

seq_name: gb_est11:AA263158

seq_documentation_block: 283 bp mRNA EST 02-JUL-1998
LOCUS AA263158
DEFINITION PM70534 KGI-a Lambda zap Express cDNA library Homo sapiens CDNA 5', mRNA sequence.
ACCESSION AA263158
VERSION AA263158.1 GI:1898964
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 283)
AUTHORS Claudio,J.O., Liew,C.C., Dempsey,A.A., Cukerman,E., Stewart,A.K., Na,E., Atkins,H.I., Iscove,N.N. and Hawley,R.G.
TITLE Identification of sequence-tagged transcripts differentially expressed within the human hematopoietic hierarchy
JOURNAL Genomics 50 (1), 44-52 (1998)
MEDLINE 98292493
COMMENT On May 5, 1995 this sequence version replaced gi:797810.
Contact: Hawley RG
Oncology Research Laboratories

```

The Toronto Hospital  
CRCS-424, 67 College St., Toronto, Ontario M5G 2M1, Canada  
Tel: 416 3403834  
Fax: 416 3403453  
Email: r.hawley@utoronto.ca  
Similar to M58636 MHC class I HLA-B\* gene. Clone was randomly  
picked from KGLA primary library.  
Seq primer: 5' GAAATTAACCTCTACTAAAGG 3'  
High quality sequence stop: 283.

## FEATURES

source  
1. 283  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="KGL-a Lambda Zap Express cDNA library"  
/cell\_type="promyeloblast"  
/cell\_line="KGL-a"  
/note="Vector: Lambda Zap Express (Stratagene); Site\_1:  
EcoRI; Site\_2: XhoI; Unidirectional cloning sites:  
EcoRI-XhoI. mRNA was purified from KGL-a cell line, cDNA  
was synthesized using an XhoI-OligodT linker primer. EcoRI  
adaptors were ligated, followed by digestion with XhoI for  
directional cloning into predigested Lambda Zap Express"

## BASE COUNT

ORIGIN 64 a 91 c 88 g 40 t

## alignment\_scores:

Quality: 44.00 Length: 10  
Ratio: 4.889 Gaps: 0  
Percent Similarity: 90.000 Percent Identity: 90.000

## alignment\_block:

US-08-653-294-25 x AA263158 ..

Align seg 1/1 to: AA263158 from: 1 to: 283

1 ArgGluAsnLeuArgThrAlaLeuArgTyr 10  
|||||  
120 CGAGAGAACCTCGGATCGCGTCCGCTAC 149

seq\_name: gb\_est6:D82221

## seq\_documentation\_block:

LOCUS D82221 375 bp mRNA EST 09-FEB-1996  
DEFINITION HUMHBC4626 Human pancreatic islet Homo sapiens cDNA similar to  
HLA-B, mRNA sequence.

ACCESSION D82221

VERSION D82221.1 GI:1183739

## KEYWORDS

EST.

## SOURCE

human.

## ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 375)

HLA-B.

mRNA sequence.

Human pancreatic islet ESTs

Unpublished (1995)

On Apr 14, 1993 this sequence version replaced gi:785255.

Contact: Jun Takeda

Institute for Molecular and Cellular Regulation, Gunma University

3-39-15 Shova-machi, Maebashi Gunma 371, Japan

Tel: 272-20-8856

Fax: 272-20-8896

Email: jtakeda@sb.gunma-u.ac.jp.

Location/Qualifiers

1. 375

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone\_lib="Human pancreatic islet"

/note="Vector: Lambda ZAPII; Site\_1: Eco RI; Site\_2: Xho

I; mRNA was prepared from normal adult human islets. cDNA

was directionally synthesized from the Xho I in the vector

to the EcoRI site. cDNA was size fractionated to remove

sequences <1000 bp in size."

BASE COUNT 75 a 124 c 118 g 55 t 3 others

## ORIGIN

## alignment\_scores:

Quality: 44.00 Length: 10  
Ratio: 4.889 Gaps: 0  
Percent Similarity: 90.000 Percent Identity: 90.000

## alignment\_block:

US-08-653-294-25 x D82221 ..

Align seg 1/1 to: D82221 from: 1 to: 375

1 ArgGluAsnLeuArgThrAlaLeuArgTyr 10

|||||

306 CGAGAGAACCTCGGATCGCGTCCGCTAC 335

seq\_name: gb\_est26:AI359260

## seq\_documentation\_block:

LOCUS AI359260 618 bp mRNA EST 15-FEB-1999

DEFINITION QY27B07.x1 NCI-CGAP\_Brn23 Homo sapiens cDNA clone IMAGE:2013205 3'

similar to gb:D32129 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,

AW-66(A-10) A\*6601 ALPHA (HUMAN);, mRNA sequence.

ACCESSION AI359260

VERSION AI359260.1 GI:4110881

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 618)

NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute / National Institute of Neurological

Disorders and Stroke, Brain Tumor Genome Anatomy Project

(CGAP/BTGA), Tumor Gene Index

Unpublished (1998)

Contact: Robert Strausberg, Ph.D.

Tel: (301) 496-1550

Email: Robert\_Strausberg@nih.gov

Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,

Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima

Bonaldo, Ph.D.

cDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 691 Std Error: 0.00

Seq primer: -400P from Gibco

High quality sequence stop: 458.

Location/Qualifiers

1. 618

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone\_image="IMAGE:2013205"

/clone\_lib="NCI-CGAP\_Brn23"

/tissue\_type="glioblastoma (pooled)"

/lab\_host="DH10B"

/note="Organ: brain; Vector: pT73D-Pac (Pharmacia) with a

modified polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st

strand cDNA was primed with a Not I - oligo(dT) primer [5'

TGTTACCAATCTGAGTGGCGGCATATCTTTTTTTTTTTTTTTTTTTT

T 3']; double-stranded cDNA was ligated to Eco RI

adaptors (Pharmacia), digested with Not I and cloned into

the Not I and Eco RI sites of the modified pT73 vector.

Library is normalized, and was constructed by Bento

Soares and M. Fatima Bonaldo."

BASE COUNT 128 a 171 c 182 g 137 t

## ORIGIN

alignment\_scores:  
 Quality: 44.00 Length: 10  
 Ratio: 4.889 Gaps: 0  
 Percent Similarity: 90.000 Percent Identity: 90.000

## alignment\_block:

US-08-653-294-25 x AI359260/rev ..  
 Align seg 1/1 to reverse of: AI359260 from: 1 to: 618

1 ArgGluAsnLeuArgThrAlaLeuArgTyr 10

|||||  
 319 CGAGAGACCTGCGATCGCGCTCCGCTAC 290

seq\_name: gb\_est31:AI696864

seq\_documentation\_block: 748 bp mRNA EST 03-JUN-1999  
 LOCUS AI696864  
 DEFINITION wc74h11.x1 NCI\_CGAP\_Pan1 Homo sapiens cDNA clone IMAGE:2324421 3'  
 similar to gb:M28205 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,  
 B-51(B-5) B\*5101 ALPHA (HUMAN);, mRNA sequence.

ACCESSION AI696864

VERSION AI696864.1 GI:4984764

## KEYWORDS

EST.

## SOURCE

human.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 748)

AUTHORS

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

JOURNAL

On Mar 16, 1998 this sequence version replaced gi:2961758.

COMMENT

Contact: Robert Strausberg, Ph.D.

Tel: (301) 496-1550

Email: Robert.Strausberg@nih.gov

Life Technologies catalog #: 11548-013

DNA Sequencing by: Washington University Genome Sequencing Center

Clone Distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

[www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)

Seq primer: -40UP from Gibco

High quality sequence stop: 424.

Location/Qualifiers

1..748

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:2324421"

/clone\_lib="NCI\_CGAP\_Pan1"

/tissue\_type="adenocarcinoma"

/lab\_host="DH10B"

/note="Organ: pancreas; Vector: pCMV-SPORT6; Site:1; salI;

Site:2; NotI; Cloned unidirectionally. Primer: Oligo dT.

Average insert size 1.72 kb. Life Technologies catalog #:

11548-013"

BASE COUNT 169 a 227 c 237 g 108 t 7 others

ORIGIN

## alignment\_scores:

Quality: 44.00 Length: 10  
 Ratio: 4.889 Gaps: 0  
 Percent Similarity: 90.000 Percent Identity: 90.000

## alignment\_block:

US-08-653-294-25 x AI696864 ..

Align seg 1/1 to: AI696864 from: 1 to: 748

1 ArgGluAsnLeuArgThrAlaLeuArgTyr 10  
 |||||  
 137 CGAGAGACCTGCGATCGCGCTCCGCTAC 166

seq\_name: gb\_est27:AI466429

seq\_documentation\_block:

LOCUS AI466429 366 bp mRNA EST 09-MAR-1999

DEFINITION vx35b04.y1 Stratagene mouse lung 937302 Mus musculus cDNA clone

IMAGE:1277167 5' similar to gb:L11570 HLA CLASS I

HISTOCOMPATIBILITY ANTIGEN, BW-72(BW-70) B\*1503 (HUMAN); gb:M69073

Mus musculus mRNA, complete cds (MOUSE);, mRNA sequence.

ACCESSION AI466429

VERSION AI466429.1 GI:4315048

KEYWORDS

EST.

SOURCE

house mouse.

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 366)

AUTHORS

Marras, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,

Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, F.,

Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R.,

Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R.,

Waterston, R. and Willson, R.

The WashU-NCI Mouse EST Project 1999

Unpublished (1999)

JOURNAL

On Feb 17, 1998 this sequence version replaced gi:2898763.

COMMENT

Contact: Marra M/WashU-NCI Mouse EST Project 1999

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.

MGI:668967

This read is a RESEQUENCE of a previously sequenced mouse clone

correct orientation)

Seq primer: -40RP from Gibco

High quality sequence stop: 243.

Location/Qualifiers

1..366

/organism="Mus musculus"

/strain="C57BL/6 x CBA"

/db\_xref="taxon:10090"

/clone="IMAGE:1277167"

/clone\_lib="Stratagene mouse lung 937302"

/sex="female"

/tissue\_type="lung"

/dev\_stage="6-8 month old"

/lab\_host="SOLR (kanamycin resistant)"

/note="Organ: lung; Vector: pBluescript SK-; Site:1;

ECORI; Site:2; XhoI; Cloned unidirectionally. Primer:

Oligo dT. 6-8 month old female lung and 1.5 year old male

lung were source of mRNA. Average insert size: 1.5 kb;

Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGCGACGAG

3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'.

BASE COUNT 65 a 115 c 124 g 62 t

ORIGIN

## alignment\_scores:

Quality: 43.00 Length: 10  
 Ratio: 4.778 Gaps: 0  
 Percent Similarity: 90.000 Percent Identity: 90.000

## alignment\_block:

US-08-653-294-25 x AI466429 ..

Align seg 1/1 to: AI466429 from: 1 to: 366





Generation and analysis of 280,000 human expressed sequence tags  
Genome Res. 6 (9), 807-828 (1996)

On Sep 12, 1996 this sequence version replaced gi:1393699.  
97044478

CONTACT  
Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@wustl.edu  
This clone is available royalty-free through LLNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Seq primer: -28M3 rev2 from Amersham  
High quality sequence stop: 272.

FEATURES  
Location/Qualifiers  
1..581  
/organism="Homo sapiens"  
/db\_xref="GBA:4620889"  
/db\_xref="taxon:9606"  
/clone="IMAGE:588387"  
/clone\_lib="Stratagene colon (#937204)"  
/lab\_host="SOLR cells (kanamycin resistant)"  
/note="Organ: colon; Vector: pBluescript SK-; Site\_1:  
ECORI; Site\_2: XhoI; Cloned unidirectionally. Primer:  
Oligo dr. T-84 colonic epithelial cell line. Average  
insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor  
sequence: 5' GAATTCGCACGAC 3' -3' adaptor sequence: 5'  
CTCGAGTCTTTTCTTTTCTTTT 3"

BASE COUNT 134 a 162 c 185 g 85 t 15 others  
ORIGIN

alignment\_scores:  
Quality: 40.00 Length: 10  
Ratio: 4.444 Gaps: 0  
Percent Similarity: 90.000 Percent Identity: 80.000

alignment\_block:  
US-08-6553-294-25 x AA147151 ..

Align seg 1/1 to: AA147151 from: 1 to: 581

1 ArgGluAsnLeuArgThrAlaLeuArgTyr 10  
|||||  
152 CGAGAGAACTGGGATCGGCTCGCTAC 181

seq\_name: gb\_est37:AI973838

seq\_documentation\_block:  
LOCUS AI973838 385 bp mRNA  
DEFINITION sdlic05.y1 Gm-c1020 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:  
Gm-c1020-1377 5', mRNA sequence.  
ACCESSION AI973838  
VERSION AI973838.1 GI:5770664  
KEYWORDS EST.  
SOURCE soybean.  
ORGANISM Glycine max  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core  
eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae;  
Glycine.  
1 (bases 1 to 385)  
Shoemaker.R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V.,  
Khanna.A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,  
Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,  
Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,  
Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,  
McCann,R., Waterston,R. and Willson,R.  
Public Soybean EST Project  
Unpublished (1999)  
On May 18, 1998 this sequence version replaced gi:3136764.  
CONTACT: Shoemaker R/Public Soybean EST Project  
Public Soybean EST Project

TITLE  
JOURNAL  
COMMENT

SOURCE ORGANISM  
human. Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 448)  
REFERENCE  
AUTHORS  
Krizman, D., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,  
Hillier, L., Lacy, M., Le, N., Lennon, G., Marra, M.,  
Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F.,  
Therling, B., White, Y., Wyllie, T., Waterston, R. and Wilson, R.  
WashU-NCI human EST project  
Unpublished (1997)  
TITLE  
JOURNAL  
COMMENT  
On Sep 19, 1997 this sequence version replaced qi:1520503.

Contact: Wilson RK  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 This clone is available royalty-free through LNL; contact the  
 IMAGE Consortium (info@image.lnl.gov) for further information.  
 Seq primer: -40m13 fwd. ET from Amersham  
 High quality sequence stop: 443.

Location/Qualifiers  
 1. 448  
 /organism="Homo sapiens"  
 /db\_xref="GDB:1304052"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:412358"  
 /clone\_lib="Soares.pineal.gland\_N3HPG"  
 /lab\_host="DH10B (ampicillin resistant)"  
 /note="Organ: pineal gland; Vector: p7T73D (Pharmacia)  
 with a modified polylinker; Site\_1: Not I; Site\_2: Eco RI;  
 1st strand cDNA was primed with a Not I - oligo(dT) primer  
 [5', TGTTACCAATCTGAAGTGGAGCGCGCCGTTTTTTTTTTTTTTT  
 3'], double-stranded cDNA was size selected, ligated to  
 Eco RI adaptors (Pharmacia), digested with Not I and  
 cloned into the Not I and Eco RI sites of a modified p7T73  
 vector (Pharmacia). Library constructed by Bento Soares  
 and M.Fatima Bonaldo."

92 a 89 c 87 g 180 t

alignment\_scores:  
 Quality: 39.00  
 Ratio: 3.900  
 Percent Similarity: 100.000 Percent Identity: 70.000  
 Length: 10  
 Gaps: 0

BASE COUNT  
 ORIGIN

```

US-08-6537-294-25 X AA/22000/rev
Align seg 1/1 to reverse of: AA722000 from: 1 to: 448

1 ArgGluAsnLeuArgThrAlaLeuArgTyr 10
|||||:|||||:|||||:|||||:|||||
303 AGAAGAAGCCTGAGAACCGCCTAAAATAT 274

seq_name: gb_est9:AA064631

seq_documentation_block:
LOCUS AA064631 452 bp mRNA EST 02-FEB-1997
DEFINITION zf72c03.s1 Soares_pineal_gland_N3HPG Homo sapiens
IMAGE:382468 3', mRNA sequence.
ACCESSION AA064631
VERSION AA064631.1 GI:1558875
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Enkaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 452)
REFERENCE
AUTHORS Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B.,
Chissoe, S., Dietrich, N., Dubucque, T., Favello, A., Gish, W.,

```

Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N.,  
Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L.,  
Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Mieg,J.,  
Trevaskis,E., Underwood,K., Wohlmann,P., Waterston,R., Wilson,R.  
and Marra,M.

Generation and analysis of 280,000 human expressed sequence tags  
Genome Res. 6 (9), 807-828 (1996)

97044478

# COMMENT

On Nov 29, 1993 this sequence version replaced gi:429994.

Contact: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: estewartson.wustl.edu

This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.

Insert Length: 633 Std Error: 0.00

Seq primer: -40M13 fwd. from Amersham

High quality sequence stop: 290.

# FEATURES

Location/Qualifiers

1..452

/organism="Homo sapiens"

/db\_xref="GDB:1290725"

/db\_xref="taxon:9606"

/clone="IMAGE:382468"

/clone\_lib="Soares\_pineal\_gland\_N3HPG"

/lab\_host="DH10B (ampicillin resistant)"

/note="Organ: pineal gland; Vector: pT73D (Pharmacia)

with a modified polylinker; Site\_1: Not I; Site\_2: Eco RI;

1st strand cDNA was primed with a Not I - oligo(dT) primer

[5' TGTTACCAATCTGAAGTGGAGCGCGCGTTTTTTTTTTTTTTT

3'], double-stranded cDNA was size selected, ligated to

Eco RI adapters (Pharmacia), digested with Not I and

cloned into the Not I and Eco RI sites of a modified pT73

vector (Pharmacia). Library constructed by Bento Soares

and M.Fatima Bonaldo."

BASE COUNT 98 a 88 c 91 g 174 t 1 others

# ORIGIN

# alignment\_scores:

Quality: 39.00 Length: 10

Ratio: 3.900

Percent Similarity: 100.000 Percent Identity: 70.000

Gaps: 0

# alignment\_block:

US-08-653-294-25 x AA064631/rev ..

Align seg 1/1 to reverse of: AA064631 from: 1 to: 452

1 ArgGluAsnLeuArgThrAlaLeuArgTyr 10

|||||:|||||:|||||:|||||:|||||

303 AGAGAAAGCCTGAGAACCGCCTAAATAT 274

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2000, 04:05:43 ; Search time 133.56 Seconds  
(without alignments)  
3.547 Million cell updates/sec

Title: US-08-653-294-26

Perfect score: 98  
Sequence: 1 YRLAIRLNERRENRLRY 20

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 188963 seqs, 23686106 residues

Total number of hits satisfying chosen parameters: 188963

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database : A\_Geneseq\_36:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	98	100.0	20	1 R92907	HLA-B2702 CTL modu
2	98	100.0	20	1 R95428	HLA-B2702 84-75-84
3	98	100.0	20	1 W33778	Immunomodulating d
4	93	94.9	20	1 R92908	HLA-B2702 CTL modu
5	93	94.9	20	1 R92908	HLA-B2702 CTL modu
6	93	94.9	20	1 W33791	Peptide B2702.84-7
7	93	94.9	20	1 W33792	Peptide B2702.84-7
8	88	89.8	20	1 R92910	HLA-B2702 CTL modu
9	88	89.8	20	1 W33793	Peptide B2702.84-7
10	85	86.7	20	1 R95430	HLA-B2702 84-75T/7
11	54.5	55.6	20	1 R92911	HLA-B2702 CTL modu
12	54.5	55.6	20	1 W33779	Immunomodulating d
13	49	50.0	20	1 R41208	Peptide fragment o
14	49	50.0	10	1 R83062	HLA-B2702 CTL modu
15	49	50.0	10	1 R95413	Alpha1-helix of HL
16	49	50.0	10	1 W07512	T-cell modulating
17	49	50.0	10	1 W47265	Immunomodulatory p
18	49	50.0	10	1 W47266	Immunomodulatory p
19	49	50.0	10	1 W33784	Peptide B2702.75-8
20	49	50.0	15	1 R92912	HLA-B2702 CTL modu
21	49	50.0	15	1 W33795	Peptide B2702.70-8
22	49	50.0	25	1 R48286	Peptide fragment o
23	49	50.0	25	1 R48286	Peptide fragment o
24	49	50.0	25	1 R83090	HLA-B2702 CTL modu
25	49	50.0	25	1 R83093	HLA-B2702 CTL modu
26	49	50.0	25	1 R95416	HLA-B2702.60-84. C
27	49	50.0	25	1 R95422	HLA-B2702.60-84. C
28	49	50.0	25	1 W33794	Peptide B2702.60-8
29	49	50.0	184	1 R06801	Peptide Seq ID No:
30	49	50.0	362	1 Y03142	Sequence of HLA-Bw
31	49	50.0	362	1 R03144	Sequence of HLA-B5
32	49	50.0	362	1 R12463	HLA-Bw53 exon. Jann
33	46	46.9	326	1 W79969	Methanococcus jann
34	44	44.9	10	1 R83094	HLA-B2702 CTL modu

35 44 44.9 10 1 R83095 HLA-B2702 CTL modu  
36 44 44.9 10 1 R83096 HLA-B2702 CTL modu  
37 44 44.9 10 1 R95425 HLA-B2702.75-84(D)  
38 44 44.9 10 1 R95426 HLA-B2702.75-84(T)  
39 44 44.9 10 1 W07513 T-cell modulating  
40 44 44.9 10 1 W47267 Immunomodulatory p  
41 44 44.9 10 1 W47268 Immunomodulatory p  
42 44 44.9 10 1 W47269 Immunomodulatory p  
43 44 44.9 10 1 W47270 Immunomodulatory p  
44 44 44.9 10 1 W33788 Peptide B2702.75-8  
45 44 44.9 10 1 W33787 Peptide B2702.75-8

#### ALIGNMENTS

RESULT 1

R92907 ID R92907 standard; peptide; 20 AA.  
AC R92907;  
DT 16-MAY-1996 (first entry)  
DE HLA-B2702 CTL modulating peptide (B2702.84-75/75-84).  
KW Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;  
KW Immunosuppressant; graft versus host disorder; transplantation; therapy;  
KW class I MHC; HLA-B2702.  
OS Synthetic.  
PN W09526979-A1.  
PD 12-OCT-1995.  
PF 03-APR-1995; U04349.  
PR 05-APR-1994; US-222851.  
PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
PI Clayberger C, Krensky AM, Parham P;  
DR WPI; 95-358582/46  
PT Extension of acceptance period of transplants from MHC unmatched  
PT donor hosts - using Class I B75-84 MHC antigen of the recipient  
PT host  
PS Example 15; Page 36; 80pp; English.  
CC R83061-R83085, R83090-R83096 and R92907-R92914 represent fragments of  
CC class I major histocompatibility complex (MHC) antigens. This sequence  
CC is an inverted dimer of residues 75-84 of the alpha-1 domain of the class  
CC I MHC HLA-B2702. These sequences can be used to extend the period of  
CC acceptance by a recipient of a transplant from an MHC unmatched donor.  
CC The peptides are administered to a patient in conjunction with a  
CC subtherapeutic amount of an immunosuppressant. This is administered to  
CC the patient for a limited period of time (compared to the lifetime  
CC administration for current treatments). The peptides particularly  
CC modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)  
CC of the patient.  
SQ Sequence 20 AA;

Query Match 100.0%; Score 98; DB 1; Length 20;  
Best Local Similarity 100.0%; Pred. No. 3.9e-10;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLAIRLNERRENRLRY 20

DB 1 YRLAIRLNERRENRLRY 20

RESULT 2

R95428 ID R95428 standard; peptide; 20 AA.  
AC R95428;  
DT 12-NOV-1996 (first entry)  
DE HLA-B2702 84-75-84 palindrome.  
KW HLA; p74; alpha1-helix; human-leucocyte-associated antigen; inhibitor;  
KW T-cell lysate; membrane protein; mammal; heat shock protein; Hsc70; APC;  
KW B cell; calcium influx; cytotoxic T lymphocyte; CTL; differentiation;  
KW cytolysis; antigen presenting cell.  
OS Synthetic.  
PN W09513288-A1.  
PD 18-MAY-1995.

PF 10-NOV-1994; U12985.  
 PR 10-NOV-1993; US-150493.  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI Clayberger C, Krensky AM;  
 DR WPI: 95-194027/25.  
 PT Compsns. comprising lymphoid surface membrane proteins - which may  
 PT inhibit cytolytic activity and differentiation of CTLs.  
 PS Example; Page 12; 29pp; English.  
 CC R95413, and R95415-R95431 represent palindromes and fragments of  
 CC human-leucocyte-associated antigens. This sequence represents the  
 CC HLA-B2702 84-75-84 palindromic. These sequences can be used to isolate  
 CC the protein p74 from a T-cell lysate. p74 is a T-cell surface membrane  
 CC protein associated with T-cell activation in mammalian T-cells, and is  
 CC also immunologically cross reactive with the heat shock protein Hsc70.  
 CC p74 is found in a limited number of cell types, but is particularly  
 CC expressed on B and T cells. p74 can be isolated by lysis of a suitable  
 CC cell with an amphoteric detergent, and then passed through an affinity  
 CC column containing a covalently bound HLA-B2702 palindromic peptide.  
 CC Compositions comprising the extracellular fragment of p74 combined with  
 CC HLA-B2702 60-84 (see R95416), induces calcium influx, and inhibits  
 CC cytotoxic T lymphocyte (CTL) differentiation or cytotoxicity. Candidate  
 CC compounds can be screened for their effect on the cytolytic activity of  
 CC T-cells, by combining them with the extracellular portion of p74 and  
 CC determining the amount of binding between the candidate compound and p74.  
 CC Modulation of CTL activity can be inhibited in a cellular composition  
 CC containing T-cells and antigen presenting cells (APCs), by adding to the  
 CC mix the extracellular portion of p74, in an amount sufficient to compete  
 CC with p74 for the binding of the p74 ligand.  
 SQ Sequence 20 AA;

Query Match 100.0%; Score 98; DB 1; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 3.9e-10;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YRLAIRLNRERENLRALRY 20  
 DB 1 YRLAIRLNRERENLRALRY 20

RESULT 3  
 W33778 3  
 ID W33778 standard; peptide; 20 AA.  
 AC W33778;  
 DT 19-JUN-1998 (first entry)  
 DE Immunomodulating dimer peptide #1.  
 DE Immunomodulating dimer; immunosuppressant drug; CTL activation;  
 KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;  
 KW rejection.  
 OS Synthetic.  
 OS Homo sapiens.  
 PN W09744351-A1.  
 PD 27-NOV-1997.  
 PF 22-MAY-1997; U08689.  
 PR 24-MAY-1996; US-653294.  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI Beulow R, Clayberger C, Krensky AM;  
 DR WPI: 98-086530/08.  
 PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B  
 PT alpha-1 domain, used for preventing rejection of transplants or  
 PT treating autoimmune diseases  
 PS Claim 16; Page 35; 41pp; English.  
 CC This sequence represents a specifically claimed immunomodulating  
 CC dimer peptide of the invention. A peptide-type compound or variant is  
 CC claimed which has immunomodulating activity, including the N-terminal  
 CC acylated and/or C-terminal amidated or esterified forms of up to 60  
 CC amino acids, where the peptide-type compound comprises the formula: A-B,  
 CC where A, B = (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or  
 CC V; aa77 = D, S or R; aa79 = R or G; aa80 = I or N; aa81, aa84 = a  
 CC hydrophobic or small amino acid; aa82 = R or L; aa83 = G or R; and aa  
 CC represents amino acid. The sequence in the brackets may optionally be  
 CC absent or truncated at any peptide type bond within the brackets. The  
 CC compounds comprise amino acid sequences related to a Class I HLA-B

CC alphas domain (positions 79-84). They can be used to inhibit cytotoxic  
 CC T-lymphocytes (CTL) from undesirably attacking cells in a host or in  
 CC vitro. They can also be used in combination with antigenic peptides or  
 CC proteins of interest to activate CTLs. They can also inhibit the  
 CC proliferation of T cells in response to anti-CD3. The peptide can be  
 CC used for preventing rejection of transplants or for treating autoimmune  
 CC diseases, e.g. diabetes, rheumatoid arthritis and lupus erythematosus.  
 CC The products can also be used for detection and diagnosis.

Query Match 100.0%; Score 98; DB 1; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 3.9e-10;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YRLAIRLNRERENLRALRY 20  
 DB 1 YRLAIRLNRERENLRALRY 20

RESULT 4  
 R92909 4  
 ID R92909 standard; peptide; 20 AA.  
 AC R92909;  
 DT 16-MAY-1996 (first entry)  
 DE HLA-B2702 CTL modulating peptide (B2702.84-75/75-84(T)).  
 DE Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;  
 KW immunosuppressant; graft versus host disorder; transplantation; therapy;  
 KW class I MHC; HLA-B2702.  
 OS Synthetic.  
 PN W09526979-A1.  
 PD 12-OCT-1995.  
 PF 05-APR-1995; U04349.  
 PR 05-APR-1994; US-222851.  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI Clayberger C, Krensky AM, Farham P;  
 DR WPI: 95-358582/46.  
 PT Extension of acceptance period of transplants from MHC unmatched  
 PT donor hosts - using Class I B75-84 MHC antigen of the recipient  
 PT host  
 PS Example 15; Page 36; 80pp; English.  
 CC R83061-R83085, R83090-R83096 and R92907-R92913 represent fragments of  
 CC class I major histocompatibility complex (MHC) antigens. This sequence  
 CC is an inverted dimer of residues 75-84 of the alpha-1 domain of the class  
 CC I MHC HLA-B2702. These sequences can be used to extend the period of  
 CC acceptance by a recipient of a transplant from an MHC unmatched donor.  
 CC The peptides are administered to a patient in conjunction with a  
 CC subtherapeutic amount of an immunosuppressant. This is administered to  
 CC the patient for a limited period of time (compared to the lifetime  
 CC administration for current treatments). The peptides particularly  
 CC modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)  
 CC of the patient.  
 SQ Sequence 20 AA;

Query Match 94.9%; Score 93; DB 1; Length 20;  
 Best Local Similarity 95.0%; Pred. No. 2.6e-09;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 YRLAIRLNRERENLRALRY 20  
 DB 1 YRLAIRLNRERENLRALRY 20

RESULT 5  
 R92908 5  
 ID R92908 standard; peptide; 20 AA.  
 AC R92908;  
 DT 16-MAY-1996 (first entry)  
 DE HLA-B2702 CTL modulating peptide (B2702.84-75(T)/75-84).  
 DE Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;  
 KW immunosuppressant; graft versus host disorder; transplantation; therapy;  
 KW class I MHC; HLA-B2702.

OS Synthetic.  
 PN W09526979-A1.  
 PD 12-OCT-1995.  
 PF 05-APR-1995.  
 PR 05-APR-1994; U04349.  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI Clayberger C, Krensky AM, Parham P;  
 DR WPI; 95-358582/46.  
 PT Extension of acceptance period of transplants from MHC unmatched  
 PT donor hosts - using Class I B75-84 MHC antigen of the recipient  
 PT host  
 PS Example 15; Page 36; 80pp; English.  
 CC R83061-R83085, R83090-R83096 and R92907-R92914 represent fragments of  
 CC class I major histocompatibility complex (MHC) antigens. This sequence  
 CC is an inverted dimer of residues 75-84 of the alpha-1 domain of the class  
 CC I MHC HLA-B\*2702. These sequences can be used to extend the period of  
 CC acceptance by a recipient of a transplant from an MHC unmatched donor.  
 CC The peptides are administered to a patient in conjunction with a  
 CC subtherapeutic amount of an immunosuppressant. This is administered to  
 CC the patient for a limited period of time (compared to the lifetime  
 CC administration for current treatments). The peptides particularly  
 CC modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)  
 CC of the patient.  
 SQ Sequence 20 AA;

Query Match 94.9%; Score 93; DB 1; Length 20;  
 Best Local Similarity 95.0%; Pred. No. 2.6e-09;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 YRLATRLNRRERENLRALRY 20  
 ||||| ||||| ||||| ||||| |||||  
 DB 1 YRLATRLNRRERENLRALRY 20

RESULT 6  
 W33791  
 ID W33791 standard; peptide; 20 AA.  
 AC W33791;  
 DT 19-JUN-1998 (first entry)  
 DE Peptide B2702.84-75/75-84 tested for immunomodulating activity.  
 KW Immunomodulating dimer; immunosuppressant drug; CTL activation;  
 KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;  
 KW rejection.  
 OS Synthetic.  
 OS Homo sapiens.  
 PN W09744351-A1.  
 PD 27-NOV-1997.  
 PF 22-MAY-1997; U08689.  
 PR 24-MAY-1996; US-653294.  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI Beulow R, Clayberger C, Krensky AM;  
 DR WPI; 98-086530/08.  
 PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B  
 PT alpha-1 domain, used for preventing rejection of transplants or  
 PT treating autoimmune diseases  
 PS Example 1; Page 19; 41pp; English.  
 CC Peptides W33784-98 and W33778-9 were assayed for their immunomodulating  
 CC activity. A peptide-type compound or variant is claimed which has  
 CC C-terminal amidated or esterified forms of up to 60 amino acids, where  
 CC the peptide-type compound comprises the formula: A-B, where A, B =  
 CC (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or V; aa77 =  
 CC D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84= a hydrophobic or  
 CC small amino acid; aa82 = R or L; aa83 = G or R; and aa represents amino  
 CC acid. The sequence in the brackets may optionally be absent or truncated  
 CC at any peptide type bond within the brackets. The compounds comprise  
 CC amino acid sequences related to a Class I HLA-B alpha1 domain (positions  
 CC 79-84). They can be used to inhibit cytotoxic T-lymphocytes (CTL) from  
 CC undesirably attacking cells in a host or in vitro. They can also be  
 CC used in combination with antigenic peptides or proteins of interest to  
 CC activate CTLs. They can also inhibit the proliferation of T cells in  
 CC response to anti-CD3. The peptide can be used for preventing rejection

CC of transplants or for treating autoimmune diseases, e.g. diabetes,  
 CC rheumatoid arthritis and lupus erythematosus. The products can also be  
 CC used for detection and diagnosis.  
 SQ Sequence 20 AA;

Query Match 94.9%; Score 93; DB 1; Length 20;  
 Best Local Similarity 95.0%; Pred. No. 2.6e-09;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 YRLATRLNRRERENLRALRY 20  
 ||||| ||||| ||||| ||||| |||||  
 DB 1 YRLATRLNRRERENLRALRY 20

RESULT 7  
 W33792  
 ID W33792 standard; peptide; 20 AA.  
 AC W33792;  
 DT 19-JUN-1998 (first entry)  
 DE Peptide B2702.84-75/75-84 tested for immunomodulating activity.  
 KW Immunomodulating dimer; immunosuppressant drug; CTL activation;  
 KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;  
 KW rejection.  
 OS Synthetic.  
 OS Homo sapiens.  
 PN W09744351-A1.  
 PD 27-NOV-1997.  
 PF 22-MAY-1997; U08689.  
 PR 24-MAY-1996; US-653294.  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI Beulow R, Clayberger C, Krensky AM;  
 DR WPI; 98-086530/08.  
 PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B  
 PT alpha-1 domain, used for preventing rejection of transplants or  
 PT treating autoimmune diseases  
 PS Example 1; Page 19; 41pp; English.  
 CC Peptides W33784-98 and W33778-9 were assayed for their immunomodulating  
 CC activity. A peptide-type compound or variant is claimed which has  
 CC immunomodulating activity, including the N-terminal acylated and/or  
 CC C-terminal amidated or esterified forms of up to 60 amino acids, where  
 CC the peptide-type compound comprises the formula: A-B, where A, B =  
 CC (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or V; aa77 =  
 CC D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84= a hydrophobic or  
 CC small amino acid; aa82 = R or L; aa83 = G or R; and aa represents amino  
 CC acid. The sequence in the brackets may optionally be absent or truncated  
 CC at any peptide type bond within the brackets. The compounds comprise  
 CC amino acid sequences related to a Class I HLA-B alpha1 domain (positions  
 CC 79-84). They can be used to inhibit cytotoxic T-lymphocytes (CTL) from  
 CC undesirably attacking cells in a host or in vitro. They can also be  
 CC used in combination with antigenic peptides or proteins of interest to  
 CC activate CTLs. They can also inhibit the proliferation of T cells in  
 CC response to anti-CD3. The peptide can be used for preventing rejection  
 CC of transplants or for treating autoimmune diseases, e.g. diabetes,  
 CC rheumatoid arthritis and lupus erythematosus. The products can also be  
 CC used for detection and diagnosis.  
 SQ Sequence 20 AA;

Query Match 94.9%; Score 93; DB 1; Length 20;  
 Best Local Similarity 95.0%; Pred. No. 2.6e-09;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 YRLATRLNRRERENLRALRY 20  
 ||||| ||||| ||||| ||||| |||||  
 DB 1 YRLATRLNRRERENLRALRY 20

RESULT 8  
 R92910  
 ID R92910 standard; peptide; 20 AA.  
 AC R92910;  
 DT 16-MAY-1996 (first entry)





R92911  
 ID R92911 standard; peptide; 20 AA.  
 AC R92911;  
 DT 16-MAY-1996 (first entry)  
 DE HLA-B2702 CTL modulating peptide (B2702.84-75/84-75).  
 KW Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;  
 KW immunosuppressant; graft versus host disorder; transplantation; therapy;  
 KW class I MHC; HLA-B2702.  
 OS Synthetic.  
 PN W09526979-A1.  
 PD 12-OCT-1995.  
 PF 05-APR-1995; U04349.  
 PR 05-APR-1994; US-222851.  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI Clayberger C, Krensky AM, Parham P;  
 DR WPI: 95-358582/46.  
 PT Extension of acceptance period of transplants from MHC unmatched  
 PT donor hosts - using Class I B75-84 MHC antigen of the recipient  
 PT host  
 PS Example 15; Page 36; 80pp; English.  
 CC R83061-R83085, R83090-R83096 and R92907-R92914 represent fragments of  
 CC class I major histocompatibility complex (MHC) antigens. This sequence  
 CC is a dimer of residues 75-84 of the alpha-1 domain of the class I MHC  
 CC HLA-B2702. These sequences can be used to extend the period of  
 CC acceptance by a recipient of a transplant from an MHC unmatched donor.  
 CC The peptides are administered to a patient in conjunction with a  
 CC subtherapeutic amount of an immunosuppressant. This is administered to  
 CC the patient for a limited period of time (compared to the lifetime  
 CC administration for current treatments). The peptides particularly  
 CC modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)  
 CC of the patient.  
 SQ Sequence 20 AA;

Query Match 55.6%; Score 54.5; DB 1; Length 20;  
 Best Local Similarity 68.4%; Pred. No. 0.0058;  
 Matches 13; Conservative 2; Mismatches 1; Indels 3; Gaps 1;

QY 1 YRLAIRLNRERLRLALR 19  
 |||||  
 Db 1 YRLAIRLNRERLRLALR 16

RESULT 12  
 W33779  
 ID W33779 standard; peptide; 20 AA.  
 AC W33779;  
 DT 19-JUN-1998 (first entry)  
 DE Immunomodulating dimer peptide #2.  
 KW Immunomodulating dimer; immunosuppressant drug; CTL activation;  
 KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;  
 KW rejection.  
 OS Synthetic.  
 OS Homo sapiens.  
 PN W09744351-A1.  
 PD 27-NOV-1997.  
 PF 22-MAY-1997; U08689.  
 PR 24-MAY-1996; US-653294.  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI Beulow R, Clayberger C, Krensky AM;  
 DR WPI: 98-086530/08.  
 PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B  
 PT alpha-1 domain, used for preventing rejection of transplants or  
 PT treating autoimmune diseases  
 PS Claim 16; Page 35; 41pp; English.  
 CC This sequence represents a specifically claimed immunomodulating  
 CC dimer peptide of the invention. A peptide-type compound or variant is  
 CC claimed which has immunomodulating activity, including the N-terminal  
 CC acylated and/or C-terminal amidated or esterified forms of up to 60  
 CC amino acids, where the peptide-type compound comprises the formula: A-B,  
 CC where A, B = (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or  
 CC V; aa77 = D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a  
 CC hydrophobic or small amino acid; aa82 = R or L; aa83 = G or R; and aa

CC represents amino acid. The sequence in the brackets may optionally be  
 CC absent or truncated at any peptide type bond within the brackets. The  
 CC compounds comprise amino acid sequences related to a Class I HLA-B  
 CC alpha domain (positions 79-84). They can be used to inhibit cytotoxic  
 CC T-lymphocytes (CTL) from undesirably attacking cells in a host or in  
 CC vitro. They can also be used in combination with antigenic peptides or  
 CC proteins of interest to activate CTLs. They can also inhibit the  
 CC proliferation of T cells in response to anti-CD3. The peptide can be  
 CC used for preventing rejection of transplants or for treating autoimmune  
 CC diseases, e.g. diabetes, rheumatoid arthritis and lupus erythematosus.  
 CC The products can also be used for detection and diagnosis.  
 SQ Sequence 20 AA;

Query Match 55.6%; Score 54.5; DB 1; Length 20;  
 Best Local Similarity 68.4%; Pred. No. 0.0058;  
 Matches 13; Conservative 2; Mismatches 1; Indels 3; Gaps 1;

QY 1 YRLAIRLNRERLRLALR 19  
 |||||  
 Db 1 YRLAIRLNRERLRLALR 16

RESULT 13  
 R41208  
 ID R41208 standard; peptide; 10 AA.  
 AC R41208;  
 DT 15-MAR-1994 (first entry)  
 DE Peptide fragment of Class I HLA peptide.  
 KW Human leukocyte antigen; HLA; peptide; transplantation; neoplasia;  
 KW parasitic disease; cytotoxic T lymphocyte; modulation.  
 OS Synthetic.  
 PN W09317699-A.  
 PD 16-SEP-1993.  
 PF 25-FEB-1993; U01758.  
 PR 02-MAR-1992; US-844716.  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI Clayberger CA, Krensky AM;  
 DR WPI: 93-303134/38.  
 PT New peptide(s) based on Class I HLA antigen domains - used for  
 PT modulating cytotoxic T-lymphocyte activity towards targets  
 PS Claim 11; Page 54; 61pp; English.  
 CC The peptide is used to modulate cytotoxic T-lymphocyte (CTL)  
 CC activity, either by inhibition or stimulation. It can be used  
 CC for inhibiting CTL toxicity in transplantations, for inducing CTL  
 CC activity in parasitic diseases and neoplasia and in studies on viral  
 CC infection. The peptide can also be used for identifying CTLs which  
 CC bind to it and removing subsets of CTLs from a T-cell composition.  
 CC This peptide sequence is more commonly found within larger peptide  
 CC compounds of not more than 30 amino acids in length.  
 SQ Sequence 10 AA;

Query Match 50.0%; Score 49; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.022;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RNLRLALRY 20  
 |||||  
 Db 1 RNLRLALRY 10

RESULT 14  
 R83062  
 ID R83062 standard; peptide; 10 AA.  
 AC R83062;  
 DT 16-MAY-1996 (first entry)  
 DE HLA-B2702 CTL modulating peptide (B2702.75-84).  
 KW Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;  
 KW immunosuppressant; graft versus host disorder; transplantation; therapy;  
 KW class I MHC; HLA-B2702.  
 OS Synthetic.  
 PN W09526979-A1.

PD 12-OCT-1995.  
 PF 05-APR-1995; U04349.  
 PR 05-APR-1994; US-222851.  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI Clayberger C, Krensky AM, Farham P;  
 DR WPI: 95-358582/46.  
 PT Extension of acceptance period of transplants from MHC unmatched  
 PT donor hosts - using Class I B75-84 MHC antigen of the recipient  
 PT host  
 PS Claim 15: Page 9: 80pp: English.  
 CC R3061-R3085, R83090-R83096 and R92907-R92913 represent fragments of  
 CC class I major histocompatibility complex (MHC) antigens. This sequence  
 CC corresponds to residues 75-84 of the alpha-1 domain of the class I MHC  
 CC HLA-B\*2702. These sequences can be used to extend the period of  
 CC acceptance by a recipient of a transplant from an MHC unmatched donor.  
 CC The peptides are administered to a patient in conjunction with a  
 CC subtherapeutic amount of an immunosuppressant. This is administered to  
 CC the patient for a limited period of time (compared to the lifetime  
 CC administration for current treatments). The peptides particularly  
 CC modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)  
 CC of the patient.  
 SQ Sequence 10 AA;

CC with p74 for the binding of the p74 ligand.  
 SQ Sequence 10 AA;

Query Match 50.0%; Score 49; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.022;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RENLRIALRY 20  
 Db 1 RENLRIALRY 10

Search completed: February 8, 2000, 04:05:43  
 Job time: 9360 sec

Query Match 50.0%; Score 49; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.022;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RENLRIALRY 20  
 Db 1 RENLRIALRY 10

RESULT 15  
 R95413  
 ID R95413 standard; peptide; 10 AA.  
 AC R95413;  
 DT 12-NOV-1996 (first entry)  
 DE Alpha-helix of HLA-B\*2702.  
 KW HLA; p74; alpha-helix; human-leucocyte-associated antigen; inhibitor;  
 KW T-cell lysate; membrane protein; mammal; heat shock protein; Hsc70; APC;  
 KW B cell; calcium influx; cytotoxic T lymphocyte; CTL; differentiation;  
 KW cytolysis; antigen presenting cell.  
 OS Synthetic.  
 PN WO9513288-A1.  
 PD 18-MAY-1995.  
 PF 10-NOV-1994; U12985.  
 PR 10-NOV-1993; US-150493.  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI Clayberger C, Krensky AM;  
 DR WPI: 95-194027/25.  
 PT Compns. comprising lymphoid surface membrane proteins - which may  
 PT inhibit cytolytic activity and differentiation of CTLs.  
 PS Example; Page 11; 29pp: English.  
 CC This sequence represents the alpha-helix of the  
 CC human-leucocyte-associated antigen B2702 (HLA-B\*2702). This sequence,  
 CC epitopes, and palindromes of it (such as R95428) can be used to isolate  
 CC the protein p74 from a T-cell lysate. p74 is a T-cell surface membrane  
 CC protein associated with T-cell activation in mammalian T-cells, and is  
 CC also immunologically cross reactive with the heat shock protein Hsc70.  
 CC p74 is found in a limited number of cell types, but is particularly  
 CC expressed on B and T cells. p74 can be isolated by lysis of a suitable  
 CC cell with an amphoteric detergent, and then passed through an affinity  
 CC column containing a covalently bound HLA-B\*2702 palindromic peptide.  
 CC Compositions comprising the extracellular fragment of p74 combined with  
 CC HLA-B\*2702.60-84 (see R95416), induces calcium influx, and inhibits  
 CC cytotoxic T lymphocyte (CTL) differentiation or cytolysis. Candidate  
 CC compounds can be screened for their effect on the cytolytic activity of  
 CC T-cells, by combining them with the extracellular portion of p74 and  
 CC determining the amount of binding between the candidate compound and p74.  
 CC Modulation of CTL activity can be inhibited in a cellular composition  
 CC containing T-cells and antigen presenting cells (APCs), by adding to the  
 CC mix the extracellular portion of p74, in an amount sufficient to compete

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 7, 2000, 18:04:37 ; Search time 111.22 Seconds  
(without alignments)  
8.482 Million cell updates/sec

Title: US-08-653-294-26

Perfect score: 98  
Sequence: 1 YRLAIRLRRENRLRY 20

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 142080 seqs, 47169319 residues

Total number of hits satisfying chosen parameters: 142080

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

PIR:62:\*  
1: Pirl:\*  
2: Pirl:\*  
3: Pirl:\*  
4: Pirl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	49	50.0	273	I38509	MHC class I histoc
2	49	50.0	274	I54463	MHC HLA-B38 chain
3	49	50.0	354	I59308	class I histocompa
4	49	50.0	354	I80168	class I histocompa
5	49	50.0	354	I80167	class I histocompa
6	49	50.0	355	I80169	class I histocompa
7	49	50.0	355	I80171	class I histocompa
8	49	50.0	359	1 HLHU12	MHC class I histoc
9	49	50.0	362	1 HLHU8	MHC class I histoc
10	49	50.0	362	B30345	MHC class I histoc
11	49	50.0	362	2 JH0341	class I histocompa
12	49	50.0	362	2 JH0339	class I histocompa
13	49	50.0	362	2 JH0340	class I histocompa
14	49	50.0	362	2 A45834	MHC class I histoc
15	49	50.0	362	2 I84486	transmembrane gly
16	49	50.0	362	2 I62045	gene HLA B-1517 pr
17	49	50.0	362	2 I84490	lymphocyte antigen
18	49	50.0	362	2 I37521	HLA-Bw57.2 antigen
19	49	50.0	362	2 A30345	MHC class I histoc
20	49	50.0	362	2 I59633	MHC HLA-B transmem
21	49	50.0	362	2 S24434	class I histocompa
22	49	50.0	362	2 I37120	MHC class I histoc
23	49	50.0	363	2 S07113	class I histocompa
24	49	50.0	363	2 S03337	class I histocompa
25	49	50.0	364	2 D35997	MHC class I histoc
26	49	50.0	365	2 S77963	MHC class I histoc
27	49	50.0	365	2 I54416	HLA-AW24 protein -
28	49	50.0	365	2 I54493	MHC class I histoc
29	44	44.9	274	1 HLHU32	MHC class I histoc
30	44	44.9	355	2 I37516	HLA-B alpha-chain

class I histocompa  
MHC class I histoc  
MHC HLA-B44.2 chain  
MHC class I histoc  
MHC class I histoc  
MHC class I histoc  
NADH dehydrogenase  
240K tegument prot  
tegument protein -  
conserved hypothet  
gene COX1 intron 4  
glucosamine-fructo  
hypothetical prote  
rRNA (adenine-N6,N  
hypothetical prote  
probable serine/th

#### ALIGNMENTS

##### RESULT 1

I38509  
MHC class I histocompatibility antigen - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 08-Sep-1996 #sequence\_revision 06-Sep-1996 #text\_change 23-Jul-1999

C:Accession: I38509

R:Cereb, N.; Choi, J.W.; Riu, K.Z.; Yang, S.Y.

Tissue Antigens 44, 271-273, 1994

A:Title: HLA-B\*5105, a newly identified B51 IEF variant.

A:Reference number: I38509; MUID:95176331

A:Accession: I38509

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-273 <RES>

A:Cross-references: EMBL:U06697; NID:g469544; PIDN:AAA92997.1; PID:g469545

C:Genetics:

A:Gene: GDB:HLA-B

A:Cross-references: GDB:120048; OMIM:142830

A:Map position: 6p21.3-6p21.3

C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 50.0%; Score 49; DB 2; Length 273;

Best Local Similarity 100.0%; Pred. No. 1.3;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RENLIRALRY 20

Db 74 RENLIRALRY 83

##### RESULT 2

I54463

MHC HLA-B38 chain - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 07-Jun-1996 #sequence\_revision 07-Jun-1996 #text\_change 23-Jul-1999

C:Accession: I54463

R:Mueller, C.A.; Engler-Blum, G.; Gekeler, V.; Steiert, I.; Weiss, E.; Schmidt, H.

Immunogenetics 30, 200-207, 1989

A:Title: Genetic and serological heterogeneity of the supertypic HLA-B locus specific

A:Reference number: I54463; MUID:89379286

A:Accession: I54463

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-274 <RES>

A:Cross-references: GB:M29864; NID:g187674; PIDN:AAA36222.1; PID:g187675

C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 50.0%; Score 49; DB 2; Length 274;

Best Local Similarity 100.0%; Pred. No. 1.3;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 RENLRIALRY 20  
| | | | | | | |  
Db 75 RENLRIALRY 84

## RESULT 3

class I histocompatibility antigen - pygmy chimpanzee (fragment)  
C:Species: Pan paniscus (pygmy chimpanzee, bonobo)  
C:Date: 31-May-1996 #sequence\_revision 31-May-1996 #text\_change 23-Jul-1999  
C:Accession: I59308  
R:McAdam, S.N.; Boyson, J.E.; Liu, X.; Garber, T.L.; Hughes, A.L.; Bontrop, R.E.; Watkin  
Proc. Natl. Acad. Sci. U.S.A. 91, 5893-5897, 1994  
A:Title: A uniquely high level of recombination at the HLA-B locus.  
A:Reference number: I59308; MUID:94286544  
A:Accession: I59308  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-354 <RES>  
A:Cross-references: EMBL:U05575; NID:g454767; PIDN:AAA50178.1; PID:g454768  
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 50.0%; Score 49; DB 2; Length 354;  
Best Local Similarity 100.0%; Pred. No. 1.7;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 RENLRIALRY 20  
| | | | | | | |  
Db 91 RENLRIALRY 100

## RESULT 4

class I histocompatibility antigen - chimpanzee (fragment)  
C:Species: Pan troglodytes (chimpanzee)  
C:Date: 24-May-1996 #sequence\_revision 24-May-1996 #text\_change 23-Jul-1999  
C:Accession: I80168  
R:McAdam, S.N.; Boyson, J.E.; Liu, X.; Garber, T.L.; Hughes, A.L.; Bontrop, R.E.; Watkin  
Proc. Natl. Acad. Sci. U.S.A. 91, 5893-5897, 1994  
A:Title: A uniquely high level of recombination at the HLA-B locus.  
A:Reference number: I59308; MUID:94286544  
A:Accession: I80168  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-354 <RES>  
A:Cross-references: EMBL:U05579; NID:g454775; PIDN:AAA50182.1; PID:g454776  
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 50.0%; Score 49; DB 2; Length 354;  
Best Local Similarity 100.0%; Pred. No. 1.7;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 RENLRIALRY 20  
| | | | | | | |  
Db 91 RENLRIALRY 100

## RESULT 5

class I histocompatibility antigen - pygmy chimpanzee (fragment)  
C:Species: Pan paniscus (pygmy chimpanzee, bonobo)  
C:Date: 24-May-1996 #sequence\_revision 24-May-1996 #text\_change 23-Jul-1999  
C:Accession: I80167  
R:McAdam, S.N.; Boyson, J.E.; Liu, X.; Garber, T.L.; Hughes, A.L.; Bontrop, R.E.; Watkin  
Proc. Natl. Acad. Sci. U.S.A. 91, 5893-5897, 1994  
A:Title: A uniquely high level of recombination at the HLA-B locus.  
A:Reference number: I59308; MUID:94286544  
A:Accession: I80167  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-354 <RES>

A:Cross-references: EMBL:U05578; NID:g454773; PIDN:AAA50181.1; PID:g454774  
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 50.0%; Score 49; DB 2; Length 354;  
Best Local Similarity 100.0%; Pred. No. 1.7;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 RENLRIALRY 20  
| | | | | | | |  
Db 91 RENLRIALRY 100

## RESULT 6

class I histocompatibility antigen - chimpanzee (fragment)  
C:Species: Pan troglodytes (chimpanzee)  
C:Date: 24-May-1996 #sequence\_revision 24-May-1996 #text\_change 23-Jul-1999  
C:Accession: I80169  
R:McAdam, S.N.; Boyson, J.E.; Liu, X.; Garber, T.L.; Hughes, A.L.; Bontrop, R.E.; Watkin  
Proc. Natl. Acad. Sci. U.S.A. 91, 5893-5897, 1994  
A:Title: A uniquely high level of recombination at the HLA-B locus.  
A:Reference number: I59308; MUID:94286544  
A:Accession: I80169  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-355 <RES>

A:Cross-references: EMBL:U05580; NID:g454777; PIDN:AAA50183.1; PID:g454778  
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 50.0%; Score 49; DB 2; Length 355;  
Best Local Similarity 100.0%; Pred. No. 1.7;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 RENLRIALRY 20  
| | | | | | | |  
Db 91 RENLRIALRY 100

## RESULT 7

class I histocompatibility antigen - chimpanzee (fragment)  
C:Species: Pan troglodytes (chimpanzee)  
C:Date: 24-May-1996 #sequence\_revision 24-May-1996 #text\_change 23-Jul-1999  
C:Accession: I80171  
R:McAdam, S.N.; Boyson, J.E.; Liu, X.; Garber, T.L.; Hughes, A.L.; Bontrop, R.E.; Watkin  
Proc. Natl. Acad. Sci. U.S.A. 91, 5893-5897, 1994  
A:Title: A uniquely high level of recombination at the HLA-B locus.  
A:Reference number: I59308; MUID:94286544  
A:Accession: I80171  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-355 <RES>

A:Cross-references: EMBL:U05582; NID:g454781; PIDN:AAA50185.1; PID:g454782  
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 50.0%; Score 49; DB 2; Length 355;  
Best Local Similarity 100.0%; Pred. No. 1.7;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 RENLRIALRY 20  
| | | | | | | |  
Db 91 RENLRIALRY 100

## RESULT 8

MHC class I histocompatibility antigen HLA alpha chain precursor (clone pHLA 12.4) -  
C:Species: Homo sapiens (man)  
C:Date: 05-Apr-1983 #sequence\_revision 05-Apr-1983 #text\_change 22-Jun-1999  
C:Accession: A02189

R:Malissen, M.; Malissen, B.; Jordan, B.R.  
Proc. Natl. Acad. Sci. U.S.A. 79, 893-897, 1982  
A:Title: Exon/intron organization and complete nucleotide sequence of an HLA gene.  
A:Reference number: A02189; MUID:82151002  
A:Accession: A02189  
A:Molecule type: DNA  
A:Residues: 1-359 <MAL>  
A:Cross-references: GB:J00181; GB:V00526; NID:g187600; PIDN:AAA36218.1; PID:g386873  
C:Comment: The seven exons correspond approximately to the domain structure of this chain  
C:Genetics:  
A:Map position: 6p21.3  
A:Introns: 22/1; 112/1; 204/1; 296/1; 335/1; 346/1  
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology  
C:Keywords: duplication; glycoprotein; heterodimer; transmembrane protein; transplantati  
F:1-21/Domain: signal sequence #status predicted <SIG>  
F:22-359/Product: class I histocompatibility antigen HLA alpha chain #status predicted <  
F:22-304/Domain: extracellular #status predicted <EXT>  
F:22-111/Domain: alpha-1 <EX1>  
F:112-203/Domain: alpha-2 <EX2>  
F:217-282/Domain: immunoglobulin homology <IMM>  
F:305-329/Domain: transmembrane #status predicted <TM>  
F:335-359/Domain: intracellular #status predicted <INT>  
F:107/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:224-280/Disulfide bonds: #status predicted

Query Match 50.0%; Score 49; DB 1; Length 359;  
Best Local Similarity 100.0%; Pred. No. 1.7;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 RENLRIALRY 20  
Db 96 RENLRIALRY 105  
|||||

## RESULT 9

HLH088  
MHC class I histocompatibility antigen HLA-B\*58 alpha chain precursor - human  
C:Species: Homo sapiens (man)  
C:Date: 31-Dec-1988 #sequence\_revision 31-Dec-1988 #text\_change 05-Sep-1997  
C:Accession: A23895  
R:Ways, J.P.; Coppin, H.L.; Parham, P.  
J. Biol. Chem. 260, 11924-11933, 1985  
A:Title: The complete primary structure of HLA-B\*58.  
A:Reference number: A23895; MUID:86008247  
A:Accession: A23895  
A:Molecule type: DNA  
A:Residues: 1-362 <WAY>  
A:Note: the authors translated the codon GCC for residue 349 as Ser  
C:Comment: This protein is a subtype of the HLA-B17 family.  
C:Genetics:  
A:Gene: GDB:HLA-B  
A:Cross-references: GDB:120048; OMTM:142830  
A:Map position: 6p21.3-6p21.3  
A:Introns: 25/1; 115/1; 207/1; 299/1; 338/1; 349/1  
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology  
C:Keywords: duplication; glycoprotein; heterodimer; transmembrane protein; transplantati  
F:1-24/Domain: signal sequence #status predicted <SIG>  
F:25-362/Product: class I histocompatibility antigen HLA-B\*58 alpha chain #status predic  
F:25-307/Domain: extracellular #status predicted <EXT>  
F:25-114/Domain: alpha-1 <EX1>  
F:115-206/Domain: alpha-2 <EX2>  
F:220-285/Domain: immunoglobulin homology <IMM>  
F:308-331/Domain: transmembrane #status predicted <TM>  
F:332-362/Domain: intracellular #status predicted <INT>  
F:110/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 50.0%; Score 49; DB 1; Length 362;  
Best Local Similarity 100.0%; Pred. No. 1.7;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 RENLRIALRY 20

Db 99 RENLRIALRY 108  
|||||  
RESULT 10  
B30345  
MHC class I histocompatibility antigen HLA-B\*52 precursor - human  
C:Species: Homo sapiens (man)  
C:Date: 29-Jan-1990 #sequence\_revision 29-Jan-1990 #text\_change 16-Feb-1997  
C:Accession: B30345  
R:Hayashi, H.; Ennis, P.D.; Ariga, H.; Salter, R.D.; Parham, P.; Kano, K.; Takiguchi, J.  
Immunol. 142, 306-311, 1989  
A:Title: HLA-B\*51 and HLA-B\*52 differ by only two amino acids which are in the helical  
A:Reference number: A30345; MUID:89080265  
A:Accession: B30345  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-362 <HAY>  
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology  
C:Keywords: transmembrane protein  
F:220-285/Domain: immunoglobulin homology <IMM>

Query Match 50.0%; Score 49; DB 2; Length 362;  
Best Local Similarity 100.0%; Pred. No. 1.7;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 RENLRIALRY 20  
Db 99 RENLRIALRY 108  
|||||

## RESULT 11

JH0541  
Class I histocompatibility antigen Gogo-B0103 heavy chain precursor - lowland gorilla  
C:Species: Gorilla gorilla gorilla (lowland gorilla)  
C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 23-Jul-1999  
C:Accession: JH0541  
R:Lawlor, D.A.; Warren, E.; Taylor, P.; Parham, P.  
J. Exp. Med. 174, 1491-1509, 1991  
A:Title: Gorilla class I major histocompatibility complex alleles: comparison to huma  
A:Reference number: JH0534; MUID:92078860  
A:Accession: JH0541  
A:Molecule type: DNA  
A:Residues: 1-362 <LAW>  
A:Cross-references: EMBL:X60254; NID:g22869; PIDN:CAA42806.1; PID:g22870  
A:Experimental source: EBV-transformed B cell  
C:Genetics:  
A:Introns: 25/1; 115/1; 207/1; 299/1; 338/1; 349/1  
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology  
C:Keywords: transmembrane protein  
F:1-24/Domain: signal sequence #status predicted <SIG>  
F:25-362/Product: class I histocompatibility antigen heavy chain, Gogo-B0103 #status  
F:25-114/Domain: alpha-1 <AL1>  
F:115-206/Domain: alpha-2 <AL2>  
F:207-298/Domain: alpha-3 <AL3>  
F:220-285/Domain: immunoglobulin homology <IMM>  
F:299-362/Domain: intracellular #status predicted <INT>

Query Match 50.0%; Score 49; DB 2; Length 362;  
Best Local Similarity 100.0%; Pred. No. 1.7;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 RENLRIALRY 20  
Db 99 RENLRIALRY 108  
|||||

## RESULT 12

JH0539  
Class I histocompatibility antigen Gogo-B0101 heavy chain precursor - lowland gorilla  
C:Species: Gorilla gorilla gorilla (lowland gorilla)

C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 23-Jul-1999  
 C:Accession: JH0539  
 R:Lawlor, D.A.; Warren, E.; Taylor, P.; Parham, P.  
 J. Exp. Med. 174, 1491-1509, 1991  
 A:Title: Gorilla class I major histocompatibility complex alleles: comparison to human a  
 A:Reference number: JH0534; MUID:92078860  
 A:Accession: JH0539  
 A:Molecule type: DNA  
 A:Residues: 1-362 <LAW>  
 A:Cross-references: EMBL:X60255; NID:g22865; PIDN:CAA42807.1; PID:g22866  
 A:Experimental source: EBV-transformed B cell  
 C:Genetics:  
 A:Introns: 25/1; 115/1; 207/1; 299/1; 338/1; 349/1  
 C:Superfamily: class I histocompatibility antigen; immunoglobulin homology  
 C:Keywords: transmembrane protein  
 F:1-24/Domain: signal sequence #status predicted <SIG>  
 F:25-362/Product: class I histocompatibility antigen heavy chain, Gogo-B0101 #status pre  
 F:25-114/Domain: alpha-1 <AL1>  
 F:115-206/Domain: alpha-2 <AL2>  
 F:207-298/Domain: alpha-3 <AL3>  
 F:220-285/Domain: immunoglobulin homology <IMM>  
 F:299-362/Domain: intracellular #status predicted <INT>

Query Match 50.0%; Score 49; DB 2; Length 362;  
 Best Local Similarity 100.0%; Pred. No. 1.7;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RENLIRALRY 20  
 |||||  
 DB 99 RENLIRALRY 108

RESULT 13  
 JH0540  
 class I histocompatibility antigen Gogo-B0102 heavy chain precursor - lowland gorilla  
 C:Species: Gorilla gorilla gorilla (lowland gorilla)  
 C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 23-Jul-1999  
 C:Accession: JH0540  
 R:Lawlor, D.A.; Warren, E.; Taylor, P.; Parham, P.  
 J. Exp. Med. 174, 1491-1509, 1991  
 A:Title: Gorilla class I major histocompatibility complex alleles: comparison to human a  
 A:Reference number: JH0534; MUID:92078860  
 A:Accession: JH0540  
 A:Molecule type: DNA  
 A:Residues: 1-362 <LAW>  
 A:Cross-references: EMBL:X60693; NID:g22867; PIDN:CAA43101.1; PID:g22868  
 A:Experimental source: EBV-transformed B cell  
 C:Genetics:  
 A:Introns: 25/1; 115/1; 207/1; 299/1; 338/1; 349/1  
 C:Superfamily: class I histocompatibility antigen; immunoglobulin homology  
 C:Keywords: transmembrane protein  
 F:1-24/Domain: signal sequence #status predicted <SIG>  
 F:25-362/Product: class I histocompatibility antigen heavy chain, Gogo-B0102 #status pre  
 F:25-114/Domain: alpha-1 <AL1>  
 F:115-206/Domain: alpha-2 <AL2>  
 F:207-298/Domain: alpha-3 <AL3>  
 F:220-285/Domain: immunoglobulin homology <IMM>  
 F:299-362/Domain: intracellular #status predicted <INT>

Query Match 50.0%; Score 49; DB 2; Length 362;  
 Best Local Similarity 100.0%; Pred. No. 1.7;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RENLIRALRY 20  
 |||||  
 DB 99 RENLIRALRY 108

RESULT 14  
 A45834  
 MHC class I histocompatibility antigen HLA-B53 alpha chain precursor - human

C:Species: Homo sapiens (man)  
 C:Date: 03-Jun-1993 #sequence\_revision 03-Jun-1993 #text\_change 23-Jul-1999  
 C:Accession: A45834  
 R:Hayashi, H.; Ooba, T.; Nakayama, S.; Sekimata, M.; Kano, K.; Takiguchi, M.  
 Immunogenetics 32, 195-199, 1990  
 A:Title: Alliospecificities between HLA-Bw53 and HLA-B35 are generated by substitution  
 A:Reference number: A45834; MUID:91033941  
 A:Accession: A45834  
 A:Molecule type: DNA  
 A:Residues: 1-362 <HAY>  
 A:Cross-references: GB:M58636; NID:g187756; PIDN:AAA36228.1; PID:g187757; GB:M33574  
 A:Note: this allele is designated B\*5301  
 C:Genetics:  
 A:Gene: GDB:HLA-B  
 A:Cross-references: GDB:120048; OMIM:142830  
 A:Map position: 6p21.3-6p21.3  
 C:Superfamily: class I histocompatibility antigen; immunoglobulin homology  
 C:Keywords: glycoprotein; heterodimer; transmembrane protein  
 F:1-24/Domain: signal sequence #status predicted <SIG>  
 F:220-285/Domain: immunoglobulin homology <IMM>  
 F:110/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match 50.0%; Score 49; DB 2; Length 362;  
 Best Local Similarity 100.0%; Pred. No. 1.7;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RENLIRALRY 20  
 |||||  
 DB 99 RENLIRALRY 108

RESULT 15  
 I84486  
 transmembrane glycoprotein - human  
 C:Species: Homo sapiens (man)  
 C:Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 23-Jul-1999  
 C:Accession: I84486  
 R:Hildebrand, W.H.; Domene, J.D.; Shen, S.Y.; Lau, M.; Terasaki, P.I.; Bunce, M.; Mar  
 Tissue Antigens 43, 209-218, 1994  
 A:Title: HLA-B\*15: a widespread and diverse family of HLA-B alleles.  
 A:Reference number: I38421; MUID:94367483  
 A:Accession: I84486  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-362 <RES>  
 A:Cross-references: GB:L15005; NID:g493154; PIDN:AAA56832.1; PID:g493155  
 C:Genetics:  
 A:Gene: HLA-B\*1513  
 C:Superfamily: class I histocompatibility antigen; immunoglobulin homology  
 C:Keywords: glycoprotein

Query Match 50.0%; Score 49; DB 2; Length 362;  
 Best Local Similarity 100.0%; Pred. No. 1.7;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RENLIRALRY 20  
 |||||  
 DB 99 RENLIRALRY 108

Search completed: February 7, 2000, 18:04:38  
 Job time: 22204 sec



3



OC Eutheria; Primates; Catarrhini; Hominidae; Gorilla.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 92078860.  
 RA LAWLOR D.A., WARREN E., TAYLOR P., PARHAM P.;  
 RT "Gorilla class I major histocompatibility complex alleles: comparison  
 to human and chimpanzee class I.";  
 RL J. Exp. Med. 174:1491-1509(1991).  
 CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO  
 THE IMMUNE SYSTEM.  
 CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
 MICROGLOBULIN).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X60254; CAA42806.1; -.  
 DR PIR: JH0541; JH0541.  
 DR HSP: P03989; IHSA.  
 DR PROSITE: PS00290; IG\_MHC; 1.  
 DR PFAM: PF00047; Ig; 1.  
 DR MHC I: Transmembrane; Glycoprotein; Signal.  
 KW SIGNAL 1 24  
 FT CHAIN 25 362  
 FT CLASS I HISTOCOMPATIBILITY ANTIGEN,  
 FT GOGO-B0103 ALPHA CHAIN.  
 FT DOMAIN 25 114  
 FT DOMAIN 115 206  
 FT DOMAIN 207 298  
 FT DOMAIN 299 308  
 FT TRANSMEM 309 332  
 FT DOMAIN 333 362  
 FT DISULFID 125 188  
 FT DISULFID 227 283  
 FT CARBOHYD 110 110  
 FT BY SIMILARITY.  
 SQ SEQUENCE 362 AA; 40248 MW; FEA6A941 CRC32;  
 Query Match 50.0%; Score 49; DB 1; Length 362;  
 Best Local Similarity 100.0%; Pred. No. 0.62;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 11 RENLIRALRY 20  
 Db |||||  
 Db 99 RENLIRALRY 108  
 RESULT 5  
 ID 1B15\_HUMAN STANDARD; PRT; 362 AA.  
 AC P10317;  
 DT 01-MAR-1989 (Rel. 10; Created)  
 DT 01-MAR-1989 (Rel. 10; Last sequence update)  
 DT 01-APR-1993 (Rel. 25; Last annotation update)  
 DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-27 B\*2702 ALPHA CHAIN  
 DE PRECURSOR (B-27K) (B27.2).  
 GN HLA-B OR HLAB.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 86220133.  
 RA SEEMANN G.H.A., REIN R.S., BROWN C.S., PLOEGH H.L.;  
 RT "Gene conversion-like mechanisms may generate polymorphism in human  
 class I genes.";  
 RL EMBO J. 5:547-552(1986).  
 RN [2]

RP SEQUENCE FROM N.A.  
 RA PARHAM P., ARNETT K.L., ADAMS E.J.;  
 RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE OF 86-107 AND 171-181.  
 RX MEDLINE: 86042671.  
 RA VEGA M.A., EQUERRA A., ROJO S., APARICIO P., BRAGADO R.,  
 RA LOPEZ DE CASTRO J.A.;  
 RT "Structural analysis of an HLA-B27 functional variant: identification  
 of residues that contribute to the specificity of recognition by  
 cytolytic T lymphocytes.";  
 RT Proc. Natl. Acad. Sci. U.S.A. 82:7394-7398(1985).  
 RL CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO  
 THE IMMUNE SYSTEM.  
 CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
 MICROGLOBULIN).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X03664; CAA27301.1; -.  
 DR EMBL: X03667; CAA27301.1; JOINED.  
 DR PIR: L38504; AAA69724.1; -.  
 DR PIR: B25092; HLHUBK.  
 DR HSP: P03989; IHSA.  
 DR MIM: I42830; -.  
 DR PROSITE: PS00290; IG\_MHC; 1.  
 DR PFAM: PF00047; Ig; 1.  
 DR MHC I: Transmembrane; Glycoprotein; Signal.  
 KW SIGNAL 1 24  
 FT CHAIN 25 362  
 FT CLASS I HISTOCOMPATIBILITY ANTIGEN,  
 FT B-27 B\*2702 ALPHA CHAIN.  
 FT DOMAIN 25 114  
 FT DOMAIN 115 206  
 FT DOMAIN 207 298  
 FT DOMAIN 299 308  
 FT TRANSMEM 309 332  
 FT DOMAIN 333 362  
 FT CARBOHYD 110 110  
 FT DISULFID 125 188  
 FT DISULFID 227 283  
 FT BY SIMILARITY.  
 SQ SEQUENCE 362 AA; 40397 MW; 9798F0BB CRC32;  
 Query Match 50.0%; Score 49; DB 1; Length 362;  
 Best Local Similarity 100.0%; Pred. No. 0.62;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 11 RENLIRALRY 20  
 Db |||||  
 Db 99 RENLIRALRY 108  
 RESULT 6  
 ID 1B47\_HUMAN STANDARD; PRT; 362 AA.  
 AC P30487;  
 DT 01-APR-1993 (Rel. 25; Created)  
 DT 01-FEB-1996 (Rel. 33; Last sequence update)  
 DT 01-FEB-1996 (Rel. 33; Last annotation update)  
 DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-49(B-21) B\*4901 ALPHA CHAIN  
 DE PRECURSOR.  
 GN HLA-B OR HLAB.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]

RP SEQUENCE FROM N.A.  
 RX MEDLINE: 89235215.  
 RA PARHAM P., LAWTON D.A., LOMEN C.E., ENNIS P.D.;  
 RT "Diversity and diversification of HLA-A,B,C alleles."  
 RL J. Immunol. 142:3937-3950(1989).  
 RN [2]  
 RP REVISION TO 78.  
 RX MEDLINE: 93056529.  
 RA HILDEBRAND W.H., MADRIGAL J.A., BELICH M.P., ZEMMOUR J., WARD F.E.,  
 RA WILLIAMS R.C., PARHAM P.;  
 RT "Serologic cross-reactivities poorly reflect allelic relationships in  
 RT the HLA-B\*12 and HLA-B\*21 groups. Dominant epitopes of the alpha 2  
 RT helix."  
 RL J. Immunol. 149:3563-3568(1992).  
 CC -1- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO  
 CC THE IMMUNE SYSTEM.  
 CC -1- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
 CC MICROGLOBULIN).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 CC EMBL: M24037; AAA02950.1;  
 DR HSSP: P30491; IALM.  
 DR MIM: 142830;  
 DR PROSITE: PS00290; IG\_MHC; 1.  
 DR PFAM: PF00047; Ig; 1.  
 DR PFAM: PF00129; MHC\_I; 1.  
 KW MHC I; Transmembrane; Glycoprotein; Signal.  
 FT SIGNAL 1 24  
 FT CHAIN 25 362  
 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,  
 B-49(B-21) B\*4901 ALPHA CHAIN.  
 FT DOMAIN 25 114  
 FT DOMAIN 115 206  
 FT DOMAIN 207 298  
 FT DOMAIN 299 309  
 FT TRANSMEM 310 333  
 FT DOMAIN 334 362  
 FT CARBOHYD 110 110  
 FT DISULFID 125 188  
 FT DISULFID 227 283  
 FT SEQUENCE 362 AA; 40581 MW; E996F82F CRC32;

Query Match 50.0%; Score 49; DB 1; Length 362;  
 Best Local Similarity 100.0%; Pred. No. 0.62;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RENLIRALRY 20  
 DB 99 RENLIRALRY 108

RESULT 7  
 1B49\_HUMAN STANDARD; PRT; 362 AA.  
 ID 1B49\_HUMAN  
 AC P18464;  
 DT 01-NOV-1990 (Rel. 16, Created)  
 DT 01-NOV-1990 (Rel. 16, Last sequence update)  
 DT 01-FEB-1996 (Rel. 33, Last annotation update)  
 DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-51(B-5) B\*5101 ALPHA CHAIN  
 DE PRECURSOR.  
 GN HLA-B OR HLAB.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.

RX MEDLINE: 90207291.  
 RA ENNIS P.D., ZEMMOUR J., SALTER R.D., PARHAM P.;  
 RT "Rapid cloning of HLA-A,B,C DNA by using the polymerase chain  
 RT reaction: frequency and nature of errors produced in amplification."  
 RL Proc. Natl. Acad. Sci. U.S.A. 87:2833-2837(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 89080265.  
 RA HAYASHI H., ENNIS P.D., ARIGA H., SALTER R.D., PARHAM P., KANO K.,  
 RA TAKIGUCHI M.;  
 RT "HLA-B\*51 and HLA-B\*52 differ by only two amino acids which are in the  
 RT helical region of the alpha 1 domain."  
 RL J. Immunol. 142:306-311(1989).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 89233295.  
 RA POHLA H., KUON W., TABACZEWSKI P., DOERNER C., WEISS E.H.;  
 RT "Allelic variation in HLA-B and HLA-C sequences and the evolution of  
 RT the HLA-B alleles."  
 RL Immunogenetics 29:297-307(1989).  
 CC -1- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO  
 CC THE IMMUNE SYSTEM.  
 CC -1- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
 CC MICROGLOBULIN).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 CC EMBL: M23219; AAA36232.1;  
 DR EMBL: M22792; AAA59620.1; ALT\_SEQ.  
 DR EMBL: M22786; AAA59620.1; JOINED.  
 DR EMBL: M22787; AAA59620.1; JOINED.  
 DR EMBL: M22788; AAA59620.1; JOINED.  
 DR EMBL: M22789; AAA59620.1; JOINED.  
 DR EMBL: M22790; AAA59620.1; JOINED.  
 DR EMBL: M22791; AAA59620.1; JOINED.  
 DR EMBL: L41087; AAA64513.1;  
 DR EMBL: L41086; AAA64513.1; JOINED.  
 DR PIR: A30345; A30345.  
 DR PIR: A30548; A30548.  
 DR HSSP: P30491; IALM.  
 DR MIM: 142830;  
 DR PROSITE: PS00290; IG\_MHC; 1.  
 DR PFAM: PF00047; Ig; 1.  
 DR PFAM: PF00129; MHC\_I; 1.  
 KW MHC I; Transmembrane; Glycoprotein; Signal.  
 FT SIGNAL 1 24  
 FT CHAIN 25 362  
 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,  
 B-51(B-5) B\*5101 ALPHA CHAIN.  
 FT DOMAIN 25 114  
 FT DOMAIN 115 206  
 FT DOMAIN 207 298  
 FT DOMAIN 299 308  
 FT TRANSMEM 309 332  
 FT DOMAIN 333 362  
 FT CARBOHYD 110 110  
 FT DISULFID 125 188  
 FT DISULFID 227 283  
 FT SEQUENCE 362 AA; 40566 MW; 4D846F30 CRC32;

Query Match 50.0%; Score 49; DB 1; Length 362;  
 Best Local Similarity 100.0%; Pred. No. 0.62;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RENLIRALRY 20  
 DB 99 RENLIRALRY 108

```
DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, BW-52(B-5) B*5201 ALPHA CHAIN
DE PRECURSOR.
GN HLA-B OR HLAB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Homiidae; Homo.
RN [1]
RX SEQUENCE FROM N.A.
RX MEDLINE; 89080265.
RA HAYASHI H., ENNIS P.D., ARIGA H., SALTER R.D., PARHAM P., KANO K.,
RA TAKIGUCHI M.;
RT "HLA-B51 and HLA-Bw52 differ by only two amino acids which are in the
RT helical region of the alpha 1 domain.";
RL J. Immunol. 142:306-311(1989).
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC THE IMMUNE SYSTEM.
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC MICROGLOBULIN).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M22799; AAA59645.1; ALT-SEQ.
CC EMBL; M22793; AAA59645.1; JOINED.
CC EMBL; M22794; AAA59645.1; JOINED.
CC EMBL; M22795; AAA59645.1; JOINED.
CC EMBL; M22796; AAA59645.1; JOINED.
CC EMBL; M22797; AAA59645.1; JOINED.
CC EMBL; M22798; AAA59645.1; JOINED.
CC PIR; B30345; B30345.
CC PIR; B30548; B30548.
CC HSSP; P30491; 1ALM.
CC MIM; 142830; -.
CC PROSITE; PS00290; IG_MHC; 1.
CC PFAM; PF00047; ig; 1.
CC PFAM; PF00129; MHC_I; 1.
CC MHC I; transmembrane; Glycoprotein; Signal.
CC SIGNAL 1 24
CC CHAIN 25 362 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
CC B-51(B-5) B*5104 ALPHA CHAIN.
CC DOMAIN 25 114 EXTRACELLULAR ALPHA-1.
CC DOMAIN 115 206 EXTRACELLULAR ALPHA-2.
CC DOMAIN 207 298 EXTRACELLULAR ALPHA-3.
CC DOMAIN 299 308 CONNECTING PEPTIDE.
CC TRANSMEM 309 332 CYTOPLASMIC TAIL.
CC DOMAIN 333 362 BY SIMILARITY.
CC CARBOHYD 110 110 BY SIMILARITY.
CC DISULFID 125 188 BY SIMILARITY.
CC DISULFID 227 283 BY SIMILARITY.
CC SEQUENCE 362 AA; 40560 MW; F22F08AB CRC32;
SQ
Query Match 50.0%; Score 49; DB 1; Length 362;
Best Local Similarity 100.0%; Pred. No. 0.62;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 11 RENLIRALRY 20
Db 99 RENLIRALRY 108
|||||
RENT 10
1B54_HUMAN STANDARD; PRT; 362 AA.
AC P30490;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-APR-1993 (Rel. 25, Last annotation update)
DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-51(B-5) B*5104 ALPHA CHAIN
DE PRECURSOR.
GN HLA-B OR HLAB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Homiidae; Homo.
RN [1]
RX SEQUENCE FROM N.A.
RX MEDLINE; 92269955.
RA BELICH M.P., MADRIGAL J.A., HILDEBRAND W.H., ZEMMOUR J.,
RA WILLIAMS R.C., LUZ R., PETZL-ERLER M.L., PARHAM P.;
RT "Unusual HLA-B alleles in two tribes of Brazilian Indians.";
RL Nature 357:326-329(1992).
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC THE IMMUNE SYSTEM.
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC MICROGLOBULIN).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; 215143; CAAT8849.1; -.
CC HSSP; P30491; 1ALM.
CC MIM; 142830; -.
CC PROSITE; PS00290; IG_MHC; 1.
CC PFAM; PF00047; ig; 1.
CC PFAM; PF00129; MHC_I; 1.
CC MHC I; transmembrane; Glycoprotein; Signal.
CC SIGNAL 1 24
CC CHAIN 25 362 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
CC B-51(B-5) B*5104 ALPHA CHAIN.
CC DOMAIN 25 114 EXTRACELLULAR ALPHA-1.
CC DOMAIN 115 206 EXTRACELLULAR ALPHA-2.
CC DOMAIN 207 298 EXTRACELLULAR ALPHA-3.
CC DOMAIN 299 308 CONNECTING PEPTIDE.
CC TRANSMEM 309 332 CYTOPLASMIC TAIL.
CC DOMAIN 333 362 BY SIMILARITY.
CC CARBOHYD 110 110 BY SIMILARITY.
CC DISULFID 125 188 BY SIMILARITY.
CC DISULFID 227 283 BY SIMILARITY.
CC SEQUENCE 362 AA; 40560 MW; F22F08AB CRC32;
SQ
Query Match 50.0%; Score 49; DB 1; Length 362;
Best Local Similarity 100.0%; Pred. No. 0.62;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 11 RENLIRALRY 20
Db 99 RENLIRALRY 108
|||||
RENT 9
1B53_HUMAN STANDARD; PRT; 362 AA.
AC P30490;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-APR-1993 (Rel. 25, Last annotation update)
```

DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, BW-53 B\*5301 ALPHA CHAIN  
DE PRECURSOR.

GN HLA-B OR HLAB.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

OC Eutheria; Primates; Catarrhini; Homiidae; Homo.

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE; 91033941.

RA HAYASHI H., Ooba T., NAKAYAMA S., SEKIMATA M., KANO K.,

RA TAKIGUCHI M.;

RT "Allospecificities between HLA-B\*53 and HLA-B\*35 are generated by

RT substitution of the residues associated with HLA-B\*4/B\*6 public

RT epitopes.";

RL Immunogenetics 32:195-199(1990).

RN [2]

RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 25-302.

RX MEDLINE; 96209672.

RA SMITH K.J., REID S.W., HARLOS K., MCMICHAEL A.J., STUART D.I.,

RA BELL J.I., JONES E.Y.;

RT "Bound water structure and polymorphic amino acids act together to

RT allow the binding of different peptides to MHC class I HLA-B\*53.";

RL Immunity 4:215-228(1996).

CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO

CC THE IMMUNE SYSTEM.

CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-

CC MICROGLOBULIN).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; M58636; AAA36228.1; -

DR PIR; A45834; A45834

DR PDB; 1A1M; 08-APR-98.

DR PDB; 1A1O; 08-APR-98.

DR MIM; 142830; -

DR PROSITE; PS00290; IG\_MHC; 1.

DR PFAM; PF00047; ig; 1.

DR PFAM; PF00129; MHC\_I; 1.

KW MHC I; Transmembrane; Glycoprotein; Signal; 3D-structure.

FT SIGNAL

FT CHAIN

FT HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,

FT BW-53 B\*5301 ALPHA CHAIN.

FT DOMAIN 25 114

FT DOMAIN 115 206

FT DOMAIN 207 298

FT DOMAIN 299 308

FT DOMAIN 309 332

FT TRANSMEM 333 362

FT DOMAIN 110 110

FT CARBOHYD 125 188

FT DISULFID 227 283

FT SEQUENCE 362 AA; 40495 MW; 2BDC746E CRC32;

SQ

Query Match 50.0%; Score 49; DB 1; Length 362;

Best Local Similarity 100.0%; Pred. No. 0.62;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 11 RENLIRALRY 20

Db 99 RENLIRALRY 108

|||||

RESULT 11

1B60\_HUMAN

ID 1B60\_HUMAN

AC P18465;

STANDARD; PRT; 362 AA.

DT

DT

DT

DE

DE

GN

OS

OC

OC

RN

RP

RX

RA

RT

RT

RL

RN

RP

RX

RA

RT

RL

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

01-NOV-1990 (Rel. 16, Created)

01-NOV-1990 (Rel. 16, Last sequence update)

01-APR-1993 (Rel. 25, Last annotation update)

HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B\*57(B-17) B\*5701 ALPHA

CHAIN PRECURSOR (B\*57.1).

HLA-B OR HLAB.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

OC Eutheria; Primates; Catarrhini; Homiidae; Homo.

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE; 90207291.

RA ENNIS P.D., ZEMMOUR J., SALTER R.D., PARHAM P.;

RT "Rapid cloning of HLA-A,B CDNA by using the polymerase chain

RT reaction; frequency and nature of errors produced in amplification.";

RL Proc. Natl. Acad. Sci. U.S.A. 87:2833-2837(1990).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE; 91067476.

RA ISAMAT M., GIRDLESTONE J., MILSTEIN C.;

RT "Nucleotide sequence of an HLA-B\*57 gene.";

RL Nucleic Acids Res. 18:6702-6702(1990).

CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO

CC THE IMMUNE SYSTEM.

CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-

CC MICROGLOBULIN).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; M32318; AAA36231.1; -

DR EMBL; X55711; CAA39244.1; -

DR PIR; S12622; S12622

DR PIR; D35997; D35997.

DR HSSP; P30491; 1A1M.

DR MIM; 142830; -

DR PROSITE; PS00290; IG\_MHC; 1.

DR PFAM; PF00047; ig; 1.

DR PFAM; PF00129; MHC\_I; 1.

KW MHC I; Transmembrane; Glycoprotein; Signal.

FT SIGNAL

FT CHAIN

FT HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,

FT B\*57(B-17) B\*5701 ALPHA CHAIN.

FT DOMAIN 25 114

FT DOMAIN 115 206

FT DOMAIN 207 298

FT DOMAIN 299 308

FT TRANSMEM 309 332

FT DOMAIN 333 362

FT CARBOHYD 110 110

FT DISULFID 125 188

FT SEQUENCE 362 AA; 40224 MW; D91DF8DD CRC32;

SQ

Query Match 50.0%; Score 49; DB 1; Length 362;

Best Local Similarity 100.0%; Pred. No. 0.62;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 11 RENLIRALRY 20

Db 99 RENLIRALRY 108

|||||

RESULT 12

1B61\_HUMAN

ID 1B61\_HUMAN

AC P30497;

STANDARD; PRT; 362 AA.

DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 01-APR-1993 (Rel. 25, Last annotation update)  
 DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-57(B-17) B\*5702 ALPHA CHAIN  
 DE PRECURSOR (BW57.2).  
 GN HLA-B OR HLAB.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 93056508.  
 RA MADRIGAL J.A., BELICH M.P., HILDEBRAND W.H., BENJAMIN R.J.,  
 RA LITTLE A.M., ZEMOUR J., ENNIS P.D., WARD F.E., PETZL-ERLER M.L.,  
 RA MARTELL R.W., DU TOIT E.D., PARHAM P.;  
 RT "Distinctive HLA-A,B antigens of black populations formed by  
 RT interallelic conversion.";  
 RL J. Immunol. 149:3411-3415(1992).  
 CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO  
 CC THE IMMUNE SYSTEM.  
 CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
 CC MICROGLOBULIN).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X61707; CAA43876.1; -.  
 DR PIR; S16774; S16774.  
 DR HSP; P30491; IALM.  
 DR MIM; 142830; -.  
 DR PROSITE; PS00290; IG\_MHC; 1.  
 DR PFAM; PF00047; ig; 1.  
 DR PFAM; PF00129; MHC.I; 1.  
 KW MHC I; Transmembrane; Glycoprotein; Signal.  
 FT SIGNAL 1 24  
 FT CHAIN 25 362  
 FT DOMAIN 25 114 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,  
 FT DOMAIN 115 206 B-57(B-17) B\*5702 ALPHA CHAIN.  
 FT DOMAIN 207 298 EXTRACELLULAR ALPHA-1.  
 FT DOMAIN 299 308 EXTRACELLULAR ALPHA-2.  
 FT TRANSMEM 309 332 EXTRACELLULAR ALPHA-3.  
 FT DOMAIN 333 362 CONNECTING PEPTIDE.  
 FT CARBOHYD 110 110 CYTOPLASMIC TAIL.  
 FT DISULFID 125 188 BY SIMILARITY.  
 FT DISULFID 227 283 BY SIMILARITY.  
 SQ SEQUENCE 362 AA; 40342 MW; 628C2156 CRC32;  
 Query Match 50.0%; Score 49; DB 1; Length 362;  
 Best Local Similarity 100.0%; Pred. No. 0.62;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 11 RENLIRALRY 20  
 Db 99 RENLIRALRY 108  
 RESULT 13  
 ID 1B62\_HUMAN STANDARD; PRT; 362 AA.  
 AC P10319;  
 DT 01-MAR-1989 (Rel. 10, Created)  
 DT 01-MAR-1989 (Rel. 10, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, BW-58(B-17) B\*5801 ALPHA  
 DE CHAIN PRECURSOR.  
 GN HLA-B OR HLAB.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.

OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 86008247.  
 RA WAYS J.P., COPPIN H.L., PARHAM P.;  
 RT "The complete primary structure of HLA-Bw58.";  
 RL J. Biol. Chem. 260:11924-11933(1985).  
 CC [2]  
 RP SEQUENCE FROM N.A.  
 RX TISSUE-BLOOD;  
 RA INQUE T., OGAWA A.;  
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO  
 CC THE IMMUNE SYSTEM.  
 CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
 CC MICROGLOBULIN).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M11799; AAA59628.1; -.  
 DR EMBL; AB008102; BAA22916.1; -.  
 DR PIR; A23895; HLHUB8.  
 DR HSP; P30491; IALM.  
 DR MIM; 142830; -.  
 DR PROSITE; PS00290; IG\_MHC; 1.  
 DR PFAM; PF00047; ig; 1.  
 DR PFAM; PF00129; MHC.I; 1.  
 KW MHC I; Transmembrane; Glycoprotein; Signal.  
 FT SIGNAL 1 24  
 FT CHAIN 25 362  
 FT DOMAIN 25 114 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,  
 FT DOMAIN 115 206 BW-58(B-17) B\*5801 ALPHA CHAIN.  
 FT DOMAIN 207 298 EXTRACELLULAR ALPHA-1.  
 FT DOMAIN 299 308 EXTRACELLULAR ALPHA-2.  
 FT TRANSMEM 309 332 EXTRACELLULAR ALPHA-3.  
 FT DOMAIN 333 362 CONNECTING PEPTIDE.  
 FT CARBOHYD 110 110 CYTOPLASMIC TAIL.  
 FT DISULFID 125 188 BY SIMILARITY.  
 FT DISULFID 227 283 BY SIMILARITY.  
 SQ SEQUENCE 362 AA; 40337 MW; 3E5E7534 CRC32;  
 Query Match 50.0%; Score 49; DB 1; Length 362;  
 Best Local Similarity 100.0%; Pred. No. 0.62;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 11 RENLIRALRY 20  
 Db 99 RENLIRALRY 108  
 RESULT 14  
 ID HLAH\_HUMAN STANDARD; PRT; 362 AA.  
 AC P01893;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 15-DEC-1999 (Rel. 39, Last annotation update)  
 DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, ALPHA CHAIN H PRECURSOR  
 DE (HLA-AR) (HLA-12.4).  
 GN HLA-H OR HLAH.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.

RN SEQUENCE FROM N.A.  
 RP MEDLINE; 82151002.  
 RA MALISSEN M., MALISSEN B., JORDAN B.R.;  
 RT "Exon/intron organization and complete nucleotide sequence of an HLA  
 gene.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 79:893-897(1982).  
 CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO  
 CC THE IMMUNE SYSTEM. COULD BE THE PRODUCT OF A PSEUDOGENE.  
 CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
 CC MICROGLOBULIN).  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; J00191; AAA36218.1; ALT\_INIT.  
 CC PIR: A02189; HLHUI2.  
 CC HSSP; P03989; IHSA.  
 CC MIM; 142800; -.  
 CC PROSITE; PS00290; IG\_MHC; 1.  
 CC PFAM; PF00047; Ig; 1.  
 CC PFAM; PF00129; MHC\_I; 1.  
 CC MHC I; Transmembrane; Glycoprotein; Signal.  
 CC SIGNAL 1 24  
 CC CHAIN 25 362 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,  
 CC ALPHA CHAIN H.  
 CC DOMAIN 25 114 EXTRACELLULAR ALPHA-1.  
 CC DOMAIN 115 206 EXTRACELLULAR ALPHA-2.  
 CC DOMAIN 207 298 EXTRACELLULAR ALPHA-3.  
 CC DOMAIN 299 308 CONNECTING PEPTIDE.  
 CC TRANSMEM 309 332  
 CC DOMAIN 333 362 CYTOPLASMIC TAIL.  
 CC CARBOHYD 110 110 BY SIMILARITY.  
 CC DISULFID 227 283 BY SIMILARITY.  
 CC SEQUENCE 362 AA; 40850 MW; 5E610F63 CRC32;

Query Match 50.0%; Score 49; DB 1; Length 362;  
 Best Local Similarity 100.0%; Pred. No. 0.62;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RENLIRALRY 20  
 DB 99 RENLIRALRY 108

RESULT 15  
 1A23\_HUMAN  
 ID 1A23\_HUMAN STANDARD; PRT; 365 AA.  
 AC P30447;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, A-23(A-9) ALPHA CHAIN  
 DE PRECURSOR.  
 GN HLA-A OR HLA-A.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A. (A\*2301).  
 RX MEDLINE; 92104637.  
 RA LITTLE A.-M., MADRIGAL J.A., PARHAM P.;  
 RT "Molecular definition of an elusive third HLA-A9 molecule: HLA-A9.3.";  
 RL Immunogenetics 35:41-45(1992).  
 CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO  
 CC THE IMMUNE SYSTEM.  
 CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-

CC MICROGLOBULIN).  
 CC -!- POLYMORPHISM: THE ONLY ALLELE OF A-23 KNOWN IS A\*2301 WHICH IS  
 CC SHOWN HERE.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; M64742; AAA03662.1; -.  
 CC HSSP; P01892; 1AQD.  
 CC MIM; 142800; -.  
 CC PROSITE; PS00290; IG\_MHC; 1.  
 CC PFAM; PF00047; Ig; 1.  
 CC PFAM; PF00129; MHC\_I; 1.  
 CC MHC I; Transmembrane; Glycoprotein; Signal.  
 CC SIGNAL 1 24  
 CC CHAIN 25 365 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,  
 CC A-23(A-9) ALPHA CHAIN.  
 CC DOMAIN 25 114 EXTRACELLULAR ALPHA-1.  
 CC DOMAIN 115 206 EXTRACELLULAR ALPHA-2.  
 CC DOMAIN 207 298 EXTRACELLULAR ALPHA-3.  
 CC DOMAIN 299 308 CONNECTING PEPTIDE.  
 CC TRANSMEM 309 332  
 CC DOMAIN 333 365 CYTOPLASMIC TAIL.  
 CC CARBOHYD 110 110 BY SIMILARITY.  
 CC DISULFID 125 188 BY SIMILARITY.  
 CC DISULFID 227 283 BY SIMILARITY.  
 CC SEQUENCE 365 AA; 40732 MW; BIC21094 CRC32;

Query Match 50.0%; Score 49; DB 1; Length 365;  
 Best Local Similarity 100.0%; Pred. No. 0.62;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RENLIRALRY 20  
 DB 99 RENLIRALRY 108

Search completed: February 8, 2000, 01:25:59  
 Job time: 1559 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2000, 19:16:13 ; Search time 176.54 Seconds  
(without alignments)  
7.855 Million cell updates/sec

Title: US-08-653-294-26

Perfect score: 98  
Sequence: 1 YRLAIRLNERENLRALRY 20

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 225878 seqs, 69334122 residues

Total number of hits satisfying chosen parameters: 225878

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

SPTREMBL\_12.\*  
1: sp-archaea.\*  
2: sp-bacteria.\*  
3: sp-fungi.\*  
4: sp-human.\*  
5: sp-invertebrate.\*  
6: sp-mammal.\*  
7: sp-mhc.\*  
8: sp-organelle.\*  
9: sp-phage.\*  
10: sp-plant.\*  
11: sp-rodent.\*  
12: sp-virus.\*  
13: sp-vertebrate.\*  
14: sp-unclassified.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result NO.	Score	Query Match %	Length	DB ID	Description
1	49	50.0	89	7 019569	019569 homo sapien
2	49	50.0	90	7 046697	046697 gorilla gor
3	49	50.0	133	7 019189	019189 homo sapien
4	49	50.0	138	7 078209	078209 homo sapien
5	49	50.0	172	7 019770	019770 homo sapien
6	49	50.0	172	7 019774	019774 homo sapien
7	49	50.0	172	7 019775	019775 homo sapien
8	49	50.0	172	7 019780	019780 homo sapien
9	49	50.0	172	7 095364	095364 homo sapien
10	49	50.0	172	7 019771	019771 homo sapien
11	49	50.0	172	7 019772	019772 homo sapien
12	49	50.0	172	7 019773	019773 homo sapien
13	49	50.0	175	7 029694	029694 homo sapien
14	49	50.0	180	7 019607	019607 homo sapien
15	49	50.0	180	7 019608	019608 homo sapien
16	49	50.0	180	7 019609	019609 homo sapien
17	49	50.0	180	7 019610	019610 homo sapien
18	49	50.0	180	7 019611	019611 homo sapien
19	49	50.0	180	7 019612	019612 homo sapien
20	49	50.0	180	7 019613	019613 homo sapien

21	49	50.0	181	7 046703	046703 homo sapien
22	49	50.0	181	7 062917	062917 homo sapien
23	49	50.0	181	7 062892	062892 homo sapien
24	49	50.0	181	7 062899	062899 homo sapien
25	49	50.0	181	7 062920	062920 homo sapien
26	49	50.0	181	7 062922	062922 homo sapien
27	49	50.0	181	7 062923	062923 homo sapien
28	49	50.0	181	7 019623	019623 homo sapien
29	49	50.0	181	7 019747	019747 homo sapien
30	49	50.0	181	7 029667	029667 homo sapien
31	49	50.0	181	7 030198	030198 homo sapien
32	49	50.0	181	7 029708	029708 homo sapien
33	49	50.0	181	7 019631	019631 homo sapien
34	49	50.0	181	7 019769	019769 homo sapien
35	49	50.0	181	7 029724	029724 homo sapien
36	49	50.0	181	7 029910	029910 homo sapien
37	49	50.0	181	7 079559	079559 homo sapien
38	49	50.0	181	7 029679	029679 homo sapien
39	49	50.0	181	7 019521	019521 homo sapien
40	49	50.0	181	7 019597	019597 homo sapien
41	49	50.0	181	7 029909	029909 homo sapien
42	49	50.0	181	7 029701	029701 homo sapien
43	49	50.0	181	7 029841	029841 homo sapien
44	49	50.0	181	7 019354	019354 gorilla gor
45	49	50.0	181	7 029765	029765 homo sapien

## ALIGNMENTS

RESULT 1  
019569 PRELIMINARY; PRT; 89 AA.  
AC 019569;  
DT 01-JAN-1998 (TREMBLrel. 05, Created)  
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)  
DE MHC CLASS I ANTIGEN (FRAGMENT).  
GN HLA-B.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA CAO K., BURDETT L., ZHANG G., FERNANDEZ-VINA M.;  
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF017330; AAB70286.2;  
KW MHC.  
FT NON\_TER 1  
FT NON\_TER 89  
SQ SEQUENCE 89 AA; 10606 MW; 99D11089 CRC32;

Query Match 50.0%; Score 49; DB 7; Length 89;  
Best Local Similarity 100.0%; Pred. No. 1.1;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RENLRALRY 20  
DB 74 RENLRALRY 83

RESULT 2  
046697 PRELIMINARY; PRT; 90 AA.  
AC 046697;  
DT 01-JUN-1998 (TREMBLrel. 06, Created)  
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)  
DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)  
DE MHC CLASS I ANTIGEN HLA-H ORTHOLOG (FRAGMENT).  
GN HLA-H.  
OS Gorilla gorilla gorilla (Lowland gorilla).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

OC Eutheria; Primates; Catarrhini; Hominidae; Gorilla.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-SHABRA;  
 RA GRIMSLEY C., MATHER K.A., OBER C.;  
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF022172; AAC99794.1; -;  
 DR PFAM; PF00129; MHC\_I; 1.  
 KW MHC.  
 FT NON\_TER 1 1  
 FT NON\_TER 90 90  
 SQ SEQUENCE 90 AA; 10689 MW; 5E5F2495 CRC32;

Query Match 50.0%; Score 49; DB 7; Length 90;  
 Best Local Similarity 100.0%; Pred. No. 1.1;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RENLIRALRY 20  
 DB 75 RENLIRALRY 84  
 |||||

RESULT 3  
 OI9189 PRELIMINARY; PRT; 133 AA.  
 AC OI9189;  
 DT 01-JAN-1998 (TRENBLrel. 05, Created)  
 DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)  
 DT 01-NOV-1999 (TRENBLrel. 12, Last annotation update)  
 DE MHC CLASS I HISTOCOMPATIBILITY ANTIGEN-B (HLA-B-27KSH) (FRAGMENT).  
 GN HLA-B.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-LEUKOCYTE;  
 RA PETERSDORF E.;  
 RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U18659; AAB60357.1; -;  
 DR MIM; 142830; -;  
 DR PFAM; PF00129; MHC\_I; 1.  
 KW MHC I.  
 FT NON\_TER 1 1  
 FT NON\_TER 133 133  
 SQ SEQUENCE 133 AA; 15491 MW; 3A3BC802 CRC32;

Query Match 50.0%; Score 49; DB 7; Length 133;  
 Best Local Similarity 100.0%; Pred. No. 1.7;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RENLIRALRY 20  
 DB 27 RENLIRALRY 36  
 |||||

RESULT 4  
 OI9189 PRELIMINARY; PRT; 138 AA.  
 AC OI9189;  
 DT 01-NOV-1998 (TRENBLrel. 08, Created)  
 DT 01-NOV-1998 (TRENBLrel. 08, Last sequence update)  
 DT 01-MAY-1999 (TRENBLrel. 10, Last annotation update)  
 DE HUMAN LEUKOCYTE ANTIGEN PRECURSOR (FRAGMENT).  
 GN HLA-A.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 98007772.

RA LAFORET M., FROELICH N., PARISSIADIS A., BAUSINGER H., PFEIFFER B.,  
 RA TONGIO M.M.;  
 RT "an intronic mutation responsible for a low level of expression of an  
 RT HLA-A\*24 allele.";  
 RL Tissue Antigens 50:340-346(1997).  
 DR EMBL; Z72423; CAA96533.1; -;  
 DR PFAM; PF00129; MHC\_I; 1.  
 KW Signal; MHC.  
 FT SIGNAL 1 24  
 FT NON\_TER 138 138  
 SQ SEQUENCE 138 AA; 15610 MW; B8417FA0 CRC32;

Query Match 50.0%; Score 49; DB 7; Length 138;  
 Best Local Similarity 100.0%; Pred. No. 1.7;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RENLIRALRY 20  
 DB 99 RENLIRALRY 108  
 |||||

RESULT 5  
 OI9770 PRELIMINARY; PRT; 172 AA.  
 AC OI9770;  
 DT 01-JAN-1998 (TRENBLrel. 05, Created)  
 DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)  
 DT 01-NOV-1998 (TRENBLrel. 08, Last annotation update)  
 DE MHC CLASS I HLA-B (FRAGMENT).  
 GN HLA-B.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA CHANDANAYINGYONG D., SIRIKONG M., LONGTA K., SRINAK D., RUNGRONG E.,  
 RA BEJCHANDRA S., JUJI T., TOKUNAGA K.;  
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U90431; AAB50144.1; -;  
 DR EMBL; U90420; AAB50144.1; JOINED.  
 DR PFAM; PF00129; MHC\_I; 1.  
 KW MHC.  
 FT NON\_TER 1 1  
 FT NON\_TER 172 172  
 SQ SEQUENCE 172 AA; 19909 MW; CAAE5641 CRC32;

Query Match 50.0%; Score 49; DB 7; Length 172;  
 Best Local Similarity 100.0%; Pred. No. 2.2;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RENLIRALRY 20  
 DB 68 RENLIRALRY 77  
 |||||

RESULT 6  
 OI9774 PRELIMINARY; PRT; 172 AA.  
 AC OI9774;  
 DT 01-JAN-1998 (TRENBLrel. 05, Created)  
 DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)  
 DT 01-NOV-1998 (TRENBLrel. 08, Last annotation update)  
 DE MHC CLASS I HLA-B (FRAGMENT).  
 GN HLA-B.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA CHANDANAYINGYONG D., SIRIKONG M., LONGTA K., SRINAK D., RUNGRONG E.,  
 RA BEJCHANDRA S., BLASCZYK R., GROSSE-WILDE H.;



RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U90423; AAB50145.1; -  
 DR EMBL; U90422; AAB50145.1; JOINED.  
 DR PFAM; PF00129; MHC\_I; 1.  
 KW MHC.  
 FT NON\_TER 1 1  
 FT NON\_TER 172 172  
 SQ SEQUENCE 172 AA; 19909 MW; CAAE5641 CRC32;

Query Match 50.0%; Score 49; DB 7; Length 172;  
 Best Local Similarity 100.0%; Pred. No. 2.2;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 RENLRIALRY 20  
 Db 68 RENLRIALRY 77

## RESULT 7

O19775 PRELIMINARY; PRT; 172 AA.  
 AC O19775;  
 DT 01-JAN-1998 (TRENBLrel. 05, Created)  
 DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)  
 DT 01-NOV-1998 (TRENBLrel. 08, Last annotation update)  
 DE MHC CLASS I HLA-B (FRAGMENT).  
 GN HLA-B.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RA CHANDANAYONG D., SIRIKONG M., LONGTA K., SRINAK D., RUNGROUNG E.,  
 RA BEJCHANDRA S.;  
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U90425; AAB50146.1; -  
 DR EMBL; U90424; AAB50146.1; JOINED.  
 DR PFAM; PF00129; MHC\_I; 1.  
 KW MHC.  
 FT NON\_TER 1 1  
 FT NON\_TER 172 172  
 SQ SEQUENCE 172 AA; 19909 MW; CAAE5641 CRC32;

Query Match 50.0%; Score 49; DB 7; Length 172;  
 Best Local Similarity 100.0%; Pred. No. 2.2;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 RENLRIALRY 20  
 Db 68 RENLRIALRY 77

## RESULT 8

O19780 PRELIMINARY; PRT; 172 AA.  
 AC O19780;  
 DT 01-JAN-1998 (TRENBLrel. 05, Created)  
 DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)  
 DT 01-NOV-1998 (TRENBLrel. 08, Last annotation update)  
 DE MHC CLASS I HLA-B (FRAGMENT).  
 GN HLA-B.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RA CHANDANAYONG D., SIRIKONG M., LONGTA K., SRINAK D., RUNGROUNG E.,  
 RA BEJCHANDRA S., JUJI T., TOKUNAGA K.;  
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U90419; AAB50143.1; -  
 DR EMBL; U90418; AAB50143.1; JOINED.

DR PFAM; PF00129; MHC\_I; 1.  
 KW MHC.  
 FT NON\_TER 1 1  
 FT NON\_TER 172 172  
 SQ SEQUENCE 172 AA; 19909 MW; CAAE5641 CRC32;

Query Match 50.0%; Score 49; DB 7; Length 172;  
 Best Local Similarity 100.0%; Pred. No. 2.2;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 RENLRIALRY 20  
 Db 68 RENLRIALRY 77

## RESULT 9

Q95364 PRELIMINARY; PRT; 172 AA.  
 ID Q95364;  
 AC Q95364;  
 DT 01-FEB-1997 (TRENBLrel. 02, Created)  
 DT 01-FEB-1997 (TRENBLrel. 02, Last sequence update)  
 DT 01-NOV-1998 (TRENBLrel. 08, Last annotation update)  
 DE MHC HLA-B\*51 PROTEIN (FRAGMENT).  
 GN HLA-B\*51FA.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA BLASCZYK R.;  
 RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; X96473; CAA65327.1; -  
 DR PFAM; PF00129; MHC\_I; 1.  
 KW MHC.  
 FT NON\_TER 1 1  
 FT NON\_TER 172 172  
 SQ SEQUENCE 172 AA; 19942 MW; 1A73E47D CRC32;

Query Match 50.0%; Score 49; DB 7; Length 172;  
 Best Local Similarity 100.0%; Pred. No. 2.2;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 RENLRIALRY 20  
 Db 65 RENLRIALRY 74

## RESULT 10

O19771 PRELIMINARY; PRT; 172 AA.  
 ID O19771;  
 AC O19771;  
 DT 01-JAN-1998 (TRENBLrel. 05, Created)  
 DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)  
 DT 01-NOV-1998 (TRENBLrel. 08, Last annotation update)  
 DE MHC CLASS I HLA-B (FRAGMENT).  
 GN HLA-B.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RA CHANDANAYONG D., SIRIKONG M., LONGTA K., SRINAK D.,  
 RA SIRIBOONRIT U., RUNGROUNG E., BEJCHANDRA S.;  
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U90612; AAB50151.1; -  
 DR EMBL; U90611; AAB50151.1; JOINED.  
 DR PFAM; PF00129; MHC\_I; 1.  
 KW MHC.  
 FT NON\_TER 1 1  
 FT NON\_TER 172 172  
 SQ SEQUENCE 172 AA; 20026 MW; 4D9A1043 CRC32;

Query Match 50.0%; Score 49; DB 7; Length 172;  
Best Local Similarity 100.0%; Pred. No. 2.2;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RENLRIALRY 20  
DB 68 RENLRIALRY 77

## RESULT 11

O19772 PRELIMINARY; PRT; 172 AA.  
AC O19772;  
DT 01-JAN-1998 (TREMBlrel. 05, Created)  
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)  
DE MHC CLASS I HLA-B (FRAGMENT).  
GN HLA-B.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA CHANDANAYINGYONG D., SIRIKONG M., LONGTA K., SRINAK D.,  
RA SIRIBOONRIT U., RUNGRONG E., BEJCHANDRA S.;  
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U90614; AAB50244.1; -;  
DR EMBL; U90613; AAB50244.1; JOINED.  
DR PFAM; PF00129; MHC\_I; 1.  
KW MHC.  
FT NON\_TER 1 1  
FT NON\_TER 172 172  
SQ SEQUENCE 172 AA; 20026 MW; 4D9A1043 CRC32;

Query Match 50.0%; Score 49; DB 7; Length 172;  
Best Local Similarity 100.0%; Pred. No. 2.2;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RENLRIALRY 20  
DB 68 RENLRIALRY 77

## RESULT 12

O19773 PRELIMINARY; PRT; 172 AA.  
AC O19773;  
DT 01-JAN-1998 (TREMBlrel. 05, Created)  
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)  
DE MHC CLASS I HLA-B (FRAGMENT).  
GN HLA-B.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA CHANDANAYINGYONG D., SIRIKONG M., LONGTA K., SRINAK D.,  
RA SIRIBOONRIT U., RUNGRONG E., BEJCHANDRA S.;  
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U90616; AAB50245.1; -;  
DR EMBL; U90615; AAB50245.1; JOINED.  
DR PFAM; PF00129; MHC\_I; 1.  
KW MHC.  
FT NON\_TER 1 1  
FT NON\_TER 172 172  
SQ SEQUENCE 172 AA; 20052 MW; F6214671 CRC32;

Query Match 50.0%; Score 49; DB 7; Length 172;

Best Local Similarity 100.0%; Pred. No. 2.2;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RENLRIALRY 20  
DB 68 RENLRIALRY 77

## RESULT 13

O29694 PRELIMINARY; PRT; 175 AA.  
AC Q29694;  
DT 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)  
DE MHC CLASS I HLA-B ANTIGEN (FRAGMENT).  
GN HLA-B.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA PETERSDORF E.;  
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U28759; AAB60367.1; -;  
DR HSSP; P10318; 1ROG.  
DR PFAM; PF00129; MHC\_I; 1.  
KW MHC.  
FT NON\_TER 1 1  
FT VARIANT 67 67 Y -> D.  
FT VARIANT 73 73 I -> T.  
FT NON\_TER 175 175  
SQ SEQUENCE 175 AA; 20332 MW; 83A0C5C3 CRC32;

Query Match 50.0%; Score 49; DB 7; Length 175;  
Best Local Similarity 100.0%; Pred. No. 2.2;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RENLRIALRY 20  
DB 68 RENLRIALRY 77

## RESULT 14

O19607 PRELIMINARY; PRT; 180 AA.  
AC O19607;  
DT 01-JAN-1998 (TREMBlrel. 05, Created)  
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)  
DE MHC CLASS I HLA-A (FRAGMENT).  
GN HLA-A.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA CHANDANAYINGYONG D., SIRIKONG M., LUANGTRAKOOL K., SRINAK D.,  
RA RUNGRONG E., BEJCHANDRA S.;  
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF030920; AAB87056.1; -;  
DR EMBL; AF030919; AAB87056.1; JOINED.  
DR HSSP; P01891; 1TMC.  
DR PFAM; PF00129; MHC\_I; 1.  
KW MHC.  
FT NON\_TER 1 1  
FT NON\_TER 180 180  
SQ SEQUENCE 180 AA; 20811 MW; CECC3537 CRC32;

Query Match 50.0%; Score 49; DB 7; Length 180;

Best Local Similarity 100.0%; Pred. No. 2.3;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RENLIRIALRY 20  
|||||  
Db 74 RENLIRIALRY 83

## RESULT 15

O19608  
ID O19608 PRELIMINARY; PRT; 180 AA.  
AC O19608;  
DT 01-JAN-1998 (TReMBLrel. 05, Created)  
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)  
DT 01-NOV-1999 (TReMBLrel. 12, Last annotation update)  
DE MHC CLASS I HLA-A (FRAGMENT).  
GN HLA-A.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA CHANDANAYONG D., SIRIKONG M., LUANGTRAKOOL K., SRINAK D.,  
RA RUNGROUNG E., BEJCHANDRA S.;  
RL Submitted (OCT-1997) to the EMBL/genBank/DBJ databases.  
DR EMBL; AF030922; AAB87057.1; -.  
DR EMBL; AF030921; AAB87057.1; JOINED.  
DR HSP; P01891; ITMC.  
DR PFAM; PF00129; MHC\_I; 1.  
KW MHC.  
FT NON\_TER 1  
FT NON\_TER 180  
SQ SEQUENCE 180 AA; 20811 MW; CECC3537 CRC32;

Query Match 50.0%; Score 49; DB 7; Length 180;

Best Local Similarity 100.0%; Pred. No. 2.3;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RENLIRIALRY 20  
|||||  
Db 74 RENLIRIALRY 83

Search completed: February 8, 2000, 19:16:13  
Job time: 21500 sec

**THIS PAGE BLANK (USPTO)**

About: Results were produced by the GenCore software, version 4.5, Copyright (c) 1993-2000 CompuGen Ltd.

seq_documentation_block:	5562 bp	DNA	INV
LOCUS	CEVITELA		
DEFINITION	C.elegans gene for vitellogenin.		

X56212gans gene for vitellogenin.  
 X56212gans  
 X56212.1 GI:6923  
 vitellogenin.  
 Caenorhabditis elegans.  
 Caenorhabditis elegans.  
 Eukaryota; Metazoa; Nematoda; Secernentea; Rhabdita; Rhabditida;  
 Rhabditina; Rhabditoidea; Rhabditidae; Pelegriniae; Caenorhabditis.  
 1 (bases 1 to 5562)  
 Blumenthal,T.  
 REFERENCE  
 AUTHORS

**TITLE** Vertebral Mollusks Conserved in Nematode and Vertebrates  
**JOURNAL** J. Mol. Evol. (1990) In press  
**REFERENCE** 2 (bases 1 to 5562)  
**AUTHORS** Blumenthal, T., Cane, J., Denison, K., Donegan, M., Kirtland, S., Lea, K., Nettleton, M., Sharrock, W., Spieth, J., Squire, M. and Zucker-Arison, E.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (05-DEC-1990) T. Blumenthal, INDIANA UNIVERSITY, DEPT OF

```

source
1:..5502
/organism="Caenorhabditis elegans"
/strain="N2, Bristol"
/db_xref="taxon:6239"
/chromosome="X"
/clone_lib="SG25, 1059 library of J. Karn"
/clone="25"
TATA_signal 437..443
prim_transcript 466..>5562
exon 466..1349
/number=1
join(466..1349,1402..1686,1734..5160,5209..5465)
mRNA
CDS
join(477..1349,1402..1686,1734..5160,5209..5465)

```

[illegible]





```

/lab_host="E.coli TGI"
/clone="CHBC1(a)"
/clone="CHBC2(A)"
/clone="CHBC4(a)"
14..250
/gene="HLA-B"
/exon
<14..>250
/note="HLA-B"
/note="allel HB(a)"
/number=2
54 a 78 c 85 g 33 t
BASE COUNT
ORIGIN

```

```

alignment_scores:
  Quality: 49.00      Length: 10
  Ratio: 4.900       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

```

```

alignment_block:
US-08-653-294-26 x HSHLABHA

```

```

Align seg 1/1 to: HSHLABHA from: 1 to: 250

```

```

11 ArgGlusnLeuArgileAlaLeuArgTyr 20
|||||
209 CGAGAGAACTGGCGATCGGCTCCGCTAC 238

```

```

seq_name: gb_prl: HSHLABHB

```

```

seq_documentation_block:
LOCUS HSHLABHB 250 bp DNA PRI
DEFINITION H.sapiens HLA-B gene, exon 2, HB(b) allele.
ACCESSION Y08693
VERSION Y08693.1 GI:1619288
KEYWORDS HLA-B gene; human leukocyte antigen.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 250)
AUTHORS Rojas-Munoz,A., Mendez,I. and Yunis,I.
TITLE Molecular evolution of HLA-B locus in a small population amerindian
community :The Nukak-Maku
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 250)
AUTHORS Rojas-Munoz,A.
TITLE Direct Submission
JOURNAL
Submitted (07-OCT-1996) A. Rojas-Munoz, National Institute Of
Health, Immunogenetics, Avenida El Dorado, Carrera 50, Santa Fe De
Bogota / Zona 6, COLOMBIA
FEATURES
source
1..250
/organism="Homo sapiens"
/isolate="Isabel-26"
/isolate="from amerindian community Nukak-Maku"
/db_xref="taxon:9606"
/chromosome="6"
/dev_stage="adult"
/tissue_type="blood"
/cell_type="white"
/lab_host="E.coli TGI"
/clone="CHBC2(b)"
/clone="CHBC3(b)"
/clone="CHBC4(b)"
14..250
/gene="HLA-B"
/exon
<14..>250
/note="HLA-B"
/note="allel HB(b)"
/number=2
51 a 78 c 87 g 34 t
BASE COUNT
ORIGIN

```

```

alignment_scores:
  Quality: 49.00      Length: 10
  Ratio: 4.900       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

```

```

alignment_block:
US-08-653-294-26 x HSHLABHB

```

```

Align seg 1/1 to: HSHLABHB from: 1 to: 250

```

```

11 ArgGlusnLeuArgileAlaLeuArgTyr 20
|||||
209 CGAGAGAACTGGCGATCGGCTCCGCTAC 238

```

```

seq_name: gb_prl: HSHLABHB

```

```

seq_documentation_block:
LOCUS HSHLABHB 250 bp DNA PRI
DEFINITION H.sapiens HLA-B gene, exon 2, HB(d) allele.
ACCESSION Y08694
VERSION Y08694.1 GI:1619289
KEYWORDS HLA-B gene; human leukocyte antigen.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 250)
AUTHORS Rojas-Munoz,A., Mendez,I. and Yunis,I.
TITLE Molecular evolution of HLA-B locus in a small population amerindian
community :The Nukak-Maku
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 250)
AUTHORS Rojas-Munoz,A.
TITLE Direct Submission
JOURNAL
Submitted (07-OCT-1996) A. Rojas-Munoz, National Institute Of
Health, Immunogenetics, Avenida El Dorado, Carrera 50, Santa Fe De
Bogota / Zona 6, COLOMBIA
FEATURES
source
1..250
/organism="Homo sapiens"
/isolate="Norman-51"
/isolate="from amerindian community Nukak-Maku"
/db_xref="taxon:9606"
/chromosome="6"
/dev_stage="adult"
/tissue_type="blood"
/cell_type="white"
/lab_host="E.coli TGI"
/clone="CHBC1(d)"
14..250
/gene="HLA-B"
/exon
<14..>250
/note="HLA-B"
/note="allel HB(d)"
/number=2
58 a 78 c 79 g 35 t
BASE COUNT
ORIGIN

```

```

alignment_scores:
  Quality: 49.00      Length: 10
  Ratio: 4.900       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

```

```

alignment_block:
US-08-653-294-26 x HSHLABHB

```

```

Align seg 1/1 to: HSHLABHB from: 1 to: 250

```

```

11 ArgGlusnLeuArgileAlaLeuArgTyr 20
|||||

```



```

209 CGAGAGAACCTGCGGATCGCGCTCCGCTAC 238
seq_name: gb_pr2:HSHLAB1
seq_documentation_block: 250 bp DNA PRI 22-MAR-1997
LOCUS HSHLAB1 Human cell line THAI DCH010 MHC class I HLA-B gene (allele
DEFINITION HLA-B*1513), exon 2.
ACCESSION U90420
VERSION U90420.1 GI:1905830
KEYWORDS
SEGMENT
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 250)
AUTHORS Chandanayingyong,D., Sirikong,M., Longta,K., Srinak,D.,
Rungroung,E., Bejchandra,S., Juji,T. and Tokunaga,K.
TITLE B15 alleles (B*1513)
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 250)
AUTHORS Chandanayingyong,D., Sirikong,M., Longta,K., Srinak,D.,
Rungroung,E., Bejchandra,S., Juji,T. and Tokunaga,K.
TITLE B15 alleles (B*1513)
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 250)
AUTHORS Chandanayingyong,D., Sirikong,M., Longta,K., Srinak,D.,
Rungroung,E., Bejchandra,S., Juji,T. and Tokunaga,K.
TITLE Direct Submission
JOURNAL Submitted (23-FEB-1997) Transfusion Medicine, Faculty of Medicine,
Siriraj Hospital, Mahidol University, Prannok Road, Bangkok 10700,
Thailand
FEATURES
Location/Qualifiers
source
1..250
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="VI"
/map="6p21"
/cell_type="lymphoblastoid"
/exon
1..250
/feature="HLA-B"
/number=2
/product="MHC class I HLA-B"
BASE COUNT 55 a 83 c 80 g 32 t
ORIGIN
11 ArgGluAsnLeuArgIleAlaLeuArgTyr 20
|||||
202 CGAGAGAACCTGCGGATCGCGCTCCGCTAC 231

alignment_scores:
Quality: 49.00 Length: 10
Ratio: 4.900 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-08-653-294-26 x HSHLAB1 ..
Align seg 1/1 to: HSHLAB1 from: 1 to: 250
11 ArgGluAsnLeuArgIleAlaLeuArgTyr 20
|||||
202 CGAGAGAACCTGCGGATCGCGCTCCGCTAC 231

seq_name: gb_pr2:HSHLAB1
seq_documentation_block: 250 bp DNA PRI 25-MAR-1997
LOCUS HSHLAB1 Human cell line THAI DCH028 MHC class I HLA-B gene (allele
DEFINITION HLA-B*51V), exon 2.
ACCESSION U90613
VERSION U90613.1 GI:1906033
KEYWORDS
SEGMENT
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 250)
AUTHORS Chandanayingyong,D., Sirikong,M., Longta,K., Srinak,D.,
Siriboonrit,U., Rungroung,E. and Bejchandra,S.
TITLE B*51V alleles
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 250)
AUTHORS Chandanayingyong,D., Sirikong,M., Longta,K., Srinak,D.,
Siriboonrit,U., Rungroung,E. and Bejchandra,S.
TITLE Direct Submission
JOURNAL Submitted (25-FEB-1997) Transfusion Medicine, Faculty of Medicine,
Siriraj Hospital, Mahidol University, Prannok Road, Bangkok 10700,
Thailand
FEATURES
Location/Qualifiers
source
1..250
/organism="Homo sapiens"

```

```

/db_xref="taxon:9606"
/chromosome="6"
/map="6p21"
/haplotype="(a)HLA-A1,B51V,Cw14,DR6(DR52),DQ1/(b)A2,B77,
Cw8N,DR4(DR53),DQ4"
/cell_type="lymphoblastoid"
/cell_line="THAI DCH028"
1..250
/gene="HLA-B"
/allele="HLA-B*51V; similar to exon 2 of B*5105"
/number=2
56 a 82 c 80 g 32 t

BASE COUNT
ORIGIN

alignment_scores:
Quality: 49.00 Length: 10
Ratio: 4.900 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-26 x HSHLABF1 ..
Align seg 1/1 to: HSHLABF1 from: 1 to: 250

11 ArgGluAsnLeuArgIleAlaLeuArgTyr 20
|||||
202 CGAGAGAACCTCGGATCGCTCGCTAC 231

seq_name: gb_pr2:HSHLABG1

seq_documentation_block:
LOCUS HSHLABG1 250 bp DNA PRI 25-MAR-1997
DEFINITION Human cell line THAI DCH011 MHC class I HLA-B gene (allele
HLA-B*51V), exon 2.
ACCESSION U90615
VERSION U90615.1 GI:1906037
KEYWORDS
SEGMENT
SOURCE
1 of 2
human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 250)
AUTHORS
Chandanayingyong,D., Sirikong,M., Longta,K., Srinak,D.,
Rungroung,E., Bejchandra,S., Blasczyk,R. and Grosse-Wilde,H.
B15 alleles (B*1513)
Unpublished
2 (bases 1 to 250)
AUTHORS
Chandanayingyong,D., Sirikong,M., Longta,K., Srinak,D.,
Rungroung,E., Bejchandra,S., Blasczyk,R. and Grosse-Wilde,H.
Direct Submission
Submitted (25-FEB-1997) Transfusion Medicine, Faculty of Medicine,
Siriraj Hospital, Mahidol University, Prannok Road, Bangkok 10700,
Thailand
FEATURES
Location/Qualifiers
1..250
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/map="6p21"
/cell_type="lymphoblastoid"
/cell_line="THAI DCH011"
1..250
/gene="HLA-B"
/allele="HLA-B*51V; similar to exon 2 of B*5105"
/number=2
56 a 82 c 80 g 32 t

BASE COUNT
ORIGIN

exon

alignment_scores:
Quality: 49.00 Length: 10
Ratio: 4.900 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-26 x HSHLABG1 ..
Align seg 1/1 to: HSHLABG1 from: 1 to: 250

11 ArgGluAsnLeuArgIleAlaLeuArgTyr 20
|||||
202 CGAGAGAACCTCGGATCGCTCGCTAC 231

seq_name: gb_pr2:HSHLABG1

seq_documentation_block:
LOCUS HSHLABG1 250 bp DNA PRI 25-MAR-1997
DEFINITION Human cell line THAI DCH011 MHC class I HLA-B gene (allele
HLA-B*51V), exon 2.
ACCESSION U90615
VERSION U90615.1 GI:1906037
KEYWORDS
SEGMENT
SOURCE
1 of 2
human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 250)
AUTHORS
Chandanayingyong,D., Sirikong,M., Longta,K., Srinak,D.,
Rungroung,E., Bejchandra,S., Blasczyk,R. and Grosse-Wilde,H.
B*51V alleles
Unpublished
2 (bases 1 to 250)
AUTHORS
Chandanayingyong,D., Sirikong,M., Longta,K., Srinak,D.,
Siriboonrit,U., Rungroung,E. and Bejchandra,S.
Direct Submission
Submitted (25-FEB-1997) Transfusion Medicine, Faculty of Medicine,
Siriraj Hospital, Mahidol University, Prannok Road, Bangkok 10700,
Thailand
FEATURES
Location/Qualifiers
1..250
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/map="6p21"
/haplotype="(a)HLA-A11.1,B51V,Cw14,DR12(DR52),DQ7/(b)A24,
B77,Cw8N,DR12(DR52),DQ7"
/cell_type="lymphoblastoid"
/cell_line="THAI DCH011"
1..250
/gene="HLA-B"
/allele="HLA-B*51V; similar to exon 2 of B*5105"
/number=2
56 a 82 c 80 g 32 t

BASE COUNT
ORIGIN

exon

```

```

alignment_scores:
Quality: 49.00 Length: 10
Ratio: 4.900 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-26 x HSHLABG1 ..
Align seg 1/1 to: HSHLABG1 from: 1 to: 250

11 ArgGluAsnLeuArgIleAlaLeuArgTyr 20
|||||
202 CGAGAGAACCTCGGATCGCTCGCTAC 231

seq_name: gb_pr2:HSHLABI1

seq_documentation_block:
LOCUS HSHLABI1 250 bp DNA PRI 22-MAR-1997
DEFINITION Human cell line THAI DCH011 MHC class I HLA-B gene (allele
HLA-B*1513), exon 2.
ACCESSION U90422
VERSION U90422.1 GI:1905834
KEYWORDS
SEGMENT
SOURCE
1 of 2
human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 250)
AUTHORS
Chandanayingyong,D., Sirikong,M., Longta,K., Srinak,D.,
Rungroung,E., Bejchandra,S., Blasczyk,R. and Grosse-Wilde,H.
B15 alleles (B*1513)
Unpublished
2 (bases 1 to 250)
AUTHORS
Chandanayingyong,D., Sirikong,M., Longta,K., Srinak,D.,
Rungroung,E., Bejchandra,S., Blasczyk,R. and Grosse-Wilde,H.
Direct Submission
Submitted (23-FEB-1997) Transfusion Medicine, Faculty of Medicine,
Siriraj Hospital, Mahidol University, Prannok Road, Bangkok 10700,
Thailand
FEATURES
Location/Qualifiers
1..250
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="VI"
/map="6p21"
/cell_type="lymphoblastoid"
/cell_line="THAI DCH011"
1..250
/gene="HLA-B"
/allele="HLA-B*1513"
/number=2
55 a 83 c 80 g 32 t

BASE COUNT
ORIGIN

alignment_scores:
Quality: 49.00 Length: 10
Ratio: 4.900 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-26 x HSHLABI1 ..
Align seg 1/1 to: HSHLABI1 from: 1 to: 250

11 ArgGluAsnLeuArgIleAlaLeuArgTyr 20
|||||
202 CGAGAGAACCTCGGATCGCTCGCTAC 231

seq_name: gb_pr2:HSHLABI1

```

```
seq_documentation_block:
LOCUS      HSHLABJ1      250 bp      DNA      PRI      22-MAR-1997
DEFINITION Human cell line THAI DCH028 MHC class I HLA-B gene (allele
ACCESSION U90424
VERSION   U90424.1 GI:1905838
KEYWORDS
SEGMENT
SOURCE
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 250)
AUTHORS   Chandanayingyong,D., Sirikong,M., Longta,K., Srinak,D.,
Rungroung,E. and Bejchandra,S.
TITLE     B15 alleles (B*1513)
JOURNAL   Unpublished
AUTHORS   Chandanayingyong,D., Sirikong,M., Longta,K., Srinak,D.,
Rungroung,E. and Bejchandra,S.
TITLE     Direct Submission
JOURNAL   Submitted (23-FEB-1997) Transfusion Medicine, Faculty of Medicine,
Siriraj Hospital, Mahidol University, Prannok Road, Bangkok 10700,
Thailand
FEATURES   Location/Qualifiers
            source
            1..250
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="VI"
            /map="6p21"
            /cell_type="lymphoblastoid"
            /cell_line="THAI DCH028"
            1..250
            /gene="HLA-B"
            /note="Allele: HLA-B*1513"
            /number=2
            /product="MHC class I HLA-B"
BASE COUNT 55 a 83 c 80 g 32 t
ORIGIN
exon
alignment_scores:
Quality: 49.00 Length: 10
Ratio: 4.900 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-08-653-294-26 x HSHLABJ1 ..
Align seg 1/1 to: HSHLABJ1 from: 1 to: 250
11 ArgGluAsnLeuArgIleAlaLeuArgTyr 20
|||||
202 CGAGAGAACCTGCGGATCGCGCTCCGCTAC 231
seq_name: gb_pr2:HSHLABJ1
seq_documentation_block:
LOCUS      HSHLABJ1      255 bp      DNA      PRI      05-JAN-1999
DEFINITION Homo sapiens isolate 026 MHC class I antigen HLA-H (HLA-H)
pseudogene, partial sequence.
ACCESSION AF022159
VERSION   AF022159.1 GI:2655062
KEYWORDS
SOURCE     human.
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 255)
AUTHORS   Grimsley,C., Mather,K.A. and Ober,C.
TITLE     HLA-H: a pseudogene with increased variation due to balancing
selection at neighboring loci
JOURNAL   Mol. Biol. Evol. 15 (12), 1581-1588 (1998)
MEDLINE   99083426
REFERENCE  2 (bases 1 to 255)
AUTHORS   Grimsley,C., Mather,K.A. and Ober,C.
TITLE     Direct Submission
JOURNAL   Submitted (03-SEP-1997) Fred Hutchinson Cancer Research Center,
1100 Fairview Ave. N., M374, Seattle, WA 98109, USA
FEATURES   Location/Qualifiers
            source
            1..255
            /organism="Homo sapiens"
            /isolate="026"
            /db_xref="taxon:9606"
            /chromosome="6"
            /map="6p21.3"
            /note="African-American individual"
            1..255
            /gene="HLA-H"
exon
seq_documentation_block:
LOCUS      HSHLABJ1      250 bp      DNA      PRI      22-MAR-1997
DEFINITION Human cell line THAI DCH009 MHC class I HLA-B gene (allele
ACCESSION U90418
VERSION   U90418.1 GI:1905826
KEYWORDS
SEGMENT
SOURCE     human.
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 250)
AUTHORS   Chandanayingyong,D., Sirikong,M., Longta,K., Srinak,D.,
Rungroung,E., Bejchandra,S., Juji,T. and Tokunaga,K.
TITLE     B15 alleles (B*1513)
```

```
/pseudo
/number=2
<1..>255
/gene="HLA-H"
/note="MHC class I antigen HLA-H"
/codon_start=1
/pseudo
<1..>255
/gene="HLA-H"
/pseudo
BASE COUNT      50 a   81 c   86 g   38 t
ORIGIN
alignment_scores:
  Quality: 49.00      Length: 10
  Ratio: 4.900       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000
alignment_block:
US-08-653-294-26 x AF022159 ..
Align seg 1/1 to: AF022159 from: 1 to: 255
11 ArgGluAsnLeuArgIleAlaLeuArgTyr 20
|||||
223 CGAGAGAACCTGCGGATCGGCTCGCTAC 252
seq_name: gb_pr4:AF022160
seq_documentation_block:
LOCUS      AF022160      259 bp      DNA
DEFINITION Homo sapiens isolate 034 MHC class I antigen HLA-H (HLA-H)
Pseudogene, partial sequence.
ACCESSION  AF022160
VERSION    AF022160.1 GI:2655063
KEYWORDS   human.
SOURCE     Homo sapiens
            Eukaryota; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 259)
AUTHORS   Grimsley,C., Mather,K.A. and Ober,C.
TITLE     HLA-H: a pseudogene with increased variation due to balancing
JOURNAL   Mol. Biol. Evol. 15 (12), 1581-1588 (1998)
MEDLINE   99083426
REFERENCE  2 (bases 1 to 259)
AUTHORS   Grimsley,C., Mather,K.A. and Ober,C.
TITLE     Direct Submission
JOURNAL   Submitted (03-SEP-1997) Fred Hutchinson Cancer Research Center,
1100 Fairview Ave. N., M374, Seattle, WA 98109, USA
FEATURES   Location/Qualifiers
            source
            1..259
            /organism="Homo sapiens"
            /isolate="034"
            /db_xref="taxon:9606"
            /chromosome="6"
            /map="6p21.3"
            /note="African-American individual"
            1..259
            /gene="HLA-H"
            /pseudo
            /number=2
            <1..>259
            /gene="HLA-H"
            /note="MHC Class I antigen HLA-H"
            /codon_start=1
            /pseudo
            <1..>259
            /gene="HLA-H"
            /pseudo
BASE COUNT      52 a   83 c   85 g   37 t   2 others
```

```
ORIGIN
alignment_scores:
  Quality: 49.00      Length: 10
  Ratio: 4.900       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000
alignment_block:
US-08-653-294-26 x AF022160 ..
Align seg 1/1 to: AF022160 from: 1 to: 259
11 ArgGluAsnLeuArgIleAlaLeuArgTyr 20
|||||
223 CGAGAGAACCTGCGGATCGGCTCGCTAC 252
seq_name: gb_pat:114590
seq_documentation_block:
LOCUS      114590      270 bp      DNA
DEFINITION Sequence 67 from patent US 5451512.
ACCESSION  114590
VERSION    114590.1 GI:997073
KEYWORDS   .
SOURCE     Unknown.
ORGANISM   Unclassified.
REFERENCE  1 (bases 1 to 270)
AUTHORS   Apple,R.J., Bugawan,T.L. and Erlich,H.A.
TITLE     Methods and reagents for HLA class I A locus DNA typing
JOURNAL   Patent: US 5451512-A 67 19-SEP-1995;
FEATURES   Location/Qualifiers
            source
            1..270
            /organism="unknown"
            55 a   84 c   95 g   36 t
BASE COUNT
ORIGIN
alignment_scores:
  Quality: 49.00      Length: 10
  Ratio: 4.900       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000
alignment_block:
US-08-653-294-26 x 114590 ..
Align seg 1/1 to: 114590 from: 1 to: 270
11 ArgGluAsnLeuArgIleAlaLeuArgTyr 20
|||||
222 CGAGAGAACCTGCGGATCGGCTCGCTAC 251
```

OM of: US-08-653-294-26 to: N\_Geneseq\_36.\* out\_format : pfs  
 Date: Feb 8, 2000 7:31 PM  
 About: Results were produced by the GenCore software, version 4.5,  
 Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:  
 -MODEL=framet\_p2n.model -DEV=xlp  
 -O=Cpnl\_1/USPTO\_spool/US08653294/runat\_04022000\_160701\_15807/app\_query.fasta.2  
 -DB=N\_Geneseq\_36 -QFMT=fastac -SUFFIX=ring -GAPOP=12.000  
 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000  
 -QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500  
 -QGAPOP=6.000 -QGAPEXT=7.000 -XGAPOP=10.000 -XGAPEXT=0.500  
 -DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blsum62  
 -TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR\_SCORE=pct  
 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0  
 -MAXLEN=1000000 -USER=US08653294 -NCPU=6 -ICPU=3 -NO\_XLPXY -WAIT  
 -THREADS=1

## Search information block:

Query: US-08-653-294-26  
 Query length: 20  
 Database: N\_Geneseq\_36.\*  
 Database sequences: 311585  
 Database length: 125096042  
 Search time (sec): 873.190000

## score\_list:

Sequence	Strid	Orig	ZScore	Escore	Len	Documentation
N_Geneseq_36:Q29167	+	49.00	153.84	0.5731	270	HLA-B*52 exon 2 alpha-1 domain
N_Geneseq_36:Q01834	+	49.00	140.24	3.28	1086	Sequence encoding HLA-B*51 anti
N_Geneseq_36:Q01822	+	49.00	140.24	3.28	1086	Sequence encoding HLA-B*52 anti
N_Geneseq_36:Q05693	+	49.00	140.22	3.29	1089	HLA-B*51 gene for production of
N_Geneseq_36:Q05701	+	49.00	140.22	3.29	1089	HLA-B*52 gene for production of
N_Geneseq_36:Q12114	+	49.00	140.22	3.29	1089	HLA-B*53 exon. HLA-B*53 gene,
N_Geneseq_36:742608	+	47.00	146.73	1.43	285	Synthetic gene, Lqhtv encoding
N_Geneseq_36:V55837	+	46.00	131.37	10.22	981	Methanococcus jannaschii FEN-1
N_Geneseq_36:V52209	-	46.00	108.69	187.51	10011	Streptococcus pneumoniae gene
N_Geneseq_36:Q28344	+	44.00	121.72	35.25	1344	pBULI core sequence from Lact
N_Geneseq_36:707447	+	43.00	131.24	10.39	362	B. subtilis aprE target DNA. De
N_Geneseq_36:Q08745	-	43.00	119.38	47.60	1220	Coding region of the aprA gene
N_Geneseq_36:Q06587	-	43.00	119.38	47.60	1220	apra gene encoding subtilin.
N_Geneseq_36:870890	-	43.00	119.38	47.60	1220	Thermotable Bacillus subtilis
N_Geneseq_36:Q03336	-	43.00	117.37	61.61	1499	Subtilisin gene. Mutant prok
N_Geneseq_36:N71241	-	43.00	117.36	61.66	1500	Subtilisin gene from Bacillus
N_Geneseq_36:N70058	-	43.00	117.36	61.66	1500	Bacillus subtilis subtilisin g
N_Geneseq_36:Q90042	-	43.00	117.36	61.66	1500	B. subtilis subtilisin gene. Re
N_Geneseq_36:N60475	-	43.00	117.20	62.90	1524	Sequence of the apr [bsu] gene
N_Geneseq_36:X25097	+	43.00	111.40	132.34	2760	Bacillus subtilis metalloprote
N_Geneseq_36:Q25826	+	42.00	138.90	3.89	118	DNA upstream from the 9th codon
N_Geneseq_36:T38732	+	42.00	138.90	3.89	118	DNA sequence upstream of 9th co
N_Geneseq_36:T19105	-	42.00	115.14	81.94	1344	Tobacco myb1 cDNA. Tobacco myb
N_Geneseq_36:T78777	+	42.00	112.40	116.41	1779	Aquifex glucosamine:fructose-6
N_Geneseq_36:T74362	+	42.00	99.41	616.20	6730	Staphylococcus aureus contig 3
N_Geneseq_36:T60544	+	41.50	126.78	18.42	345	(Pro23-Lys37)delan37-Ile-134) t
N_Geneseq_36:T60542	+	41.50	124.93	23.36	417	MetGDNF degenerate DNA sequence
N_Geneseq_36:T07008	+	41.00	137.33	4.76	99	Immunoglobulin gene from n-(ABCD) he
N_Geneseq_36:T80669	+	41.00	123.14	29.37	423	Type II topoisomerase database
N_Geneseq_36:X20619	+	41.00	113.89	96.21	1091	Polynucleotide sequence from t
N_Geneseq_36:V52187	-	41.00	85.02	3.9e+03	20986	Streptococcus pneumoniae gene
N_Geneseq_36:X13117	-	41.00	81.05	6.5e+03	31517	Enterococcus faecalis genome
N_Geneseq_36:X20057	+	40.50	112.00	122.63	1119	Enterococcus faecalis EF030 ge
N_Geneseq_36:X20056	+	40.50	110.98	139.74	1242	Enterococcus faecalis gene EF0
N_Geneseq_36:X13168	+	40.50	102.40	420.11	2991	Enterococcus faecalis genome d
N_Geneseq_36:Q13257	+	40.00	119.67	45.84	431	23S rRNA gene (partial). Nucle
N_Geneseq_36:V07462	+	40.00	119.32	47.98	447	DNA encoding the first domain.
N_Geneseq_36:Q12871	+	40.00	109.37	159.08	1164	Polyhedrin-soluble LAG-3 fusio
N_Geneseq_36:V75169	+	40.00	109.87	161.14	1176	Staphylococcus aureus contig 5
N_Geneseq_36:X06756	+	40.00	107.70	212.73	1468	Human lymphocyte activation ge
N_Geneseq_36:T60306	-	40.00	107.49	218.55	1500	Caenorhabditis elegans SEL-12

N\_Geneseq\_36:X06755 + 40.00 106.69 242.35 1629 Human lymphocyte activation  
 N\_Geneseq\_36:Q12869 + 40.00 105.33 288.25 1871 Lymphocyte Activation Gene-  
 N\_Geneseq\_36:X06754 + 40.00 103.41 369.04 2279 Human lymphocyte activation  
 N\_Geneseq\_36:Q46543 - 40.00 102.97 390.46 2384 Drosophila SSRP composite s

seq\_name: N\_Geneseq\_36:Q29167

## seq\_documentation\_block:

ID Q29167 standard; DNA; 270 BP.  
 AC Q29167;  
 DT 03-MAR-1993 (first entry)  
 DE HLA-B\*52 exon 2 alpha-1 domain.  
 KW Human leukocyte antigen; transgenic; germ cells; somatic cells;  
 KW expression; ss.  
 PN J04091731-A.  
 PD 25-MAR-1992.  
 PF 03-AUG-1990; 207329.  
 PR 03-AUG-1990; JP-207329.  
 PA (OLYU ) OLYMPUS OPTICAL CO.  
 DR WPI: 92-342893/42.  
 PT Transgenic non-human mammalian HLA-B\*52 gene - useful for  
 PT analysis of expression of gene structure, and prodn. of  
 PT mouse model of human disease  
 PS Disclosure; Fig 1; 8pp; Japanese.  
 CC The sequence shows the exon 2 alpha-1-domain of the human leukocyte  
 CC antigen-B\*52 gene. The complete gene may be introduced into non-  
 CC human mammals, pref. rat or mouse, or their ancestors at the primary  
 CC developmental biological step via transplantation into the zygote or  
 CC embryo to generate transgenic non-human mammals incorporating the  
 CC HLA-B\*52 gene in both their germ cells and somatic cells. Transgenic  
 CC non-human mammals contg. HLA-B\*52 are useful for the analysis of  
 CC expression of the gene, its structure, and prodn. of mouse models of  
 CC human disease. See also Q29166-72.  
 SQ Sequence 270 BP; 59 A; 88 C; 86 G; 37 T;

## alignment\_scores:

Quality: 49.00 Length: 10  
 Ratio: 4.900 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-08-653-294-26 x Q29167 ..

Align seg 1/1 to: Q29167 from: 1 to: 270

11 ArgGLuAsnLeuArgIleAlaLeuArgTyr 20  
 |||||  
 222 CGAGAGACCTGCGGATGCGCTCCGCTAC 251

seq\_name: N\_Geneseq\_36:Q01834

## seq\_documentation\_block:

ID Q01834 standard; DNA; 1086 BP.  
 AC Q01834;  
 DT 19-MAR-1991 (first entry)  
 DE Sequence encoding HLA-B\*51 antigen.  
 KW Probe: HLA class I DNA; immunogen; ss.  
 OS Homo sapiens.  
 PN EP35480-A.  
 PD 14-FEB-1990.  
 PF 10-AUG-1989.  
 PR 11-AUG-1988; JP-200758.  
 PA (OLYU) Olympus Optical Co., Ltd.  
 PI Kano K, Takiguchi;  
 DR WPI: 90-046289/07.  
 PT New DNA for class 1 human leucocyte antigens and derived probes and  
 PT transformed cells, useful for DNA typing, as immunogens etc.  
 PS Claim 1: Page 11; 23pp; English.  
 CC The HLA class I DNA can be used as a source of probes for use in DNA  
 CC typing. Transformed cells, which are useful as immunogens, can be  
 CC obtained by introducing these DNAs into eucaryotic cells.  
 SQ Sequence 1086 BP; 224 A; 334 C; 356 G; 172 T;

alignment\_scores:  
 Quality: 49.00 Length: 10  
 Ratio: 4.900 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-08-653-294-26 x Q01834 ..

Align seg 1/1 to: Q01834 from: 1 to: 1086

11 ArgGluAsnLeuArgIleAlaLeuArgTyr 20

|||||  
 294 CGAGAGAACCTGCGGATCGCGCTCGCTAC 323

seq\_name: N\_Geneseq\_36:Q01822

seq\_documentation\_block:

ID Q01822 standard; DNA; 1086 BP.

AC Q01822;

DT 19-MAY-1991 (first entry)

DE Sequence encoding HLA-Bw52 antigen.

KW Probe: HLA class I DNA; immunogen; ss.

OS Homo sapiens.

FH Key Location/Qualifiers

FT cds 1..1086

FT /\*tag= a

PN EP-354580-A.

PD 14-FEB-1990.

PF 10-AUG-1989.

PR 11-AUG-1988; JP-200758.

PA (OLYU) Olympos Optical Co., Ltd.

PI Kano K, Takiguchi;

DR WPI; 90-046289/07.

DR P-PSDB; R03142.

PT New DNA for class I human leucocyte antigens and derived probes and

PT transformed cells, useful for DNA typing, as immunogens etc.

PS Claim 2; pp11-12; 23pp; English.

CC The HLA class I DNA can be used as a source of probes for use in DNA

CC typing. Transformed cells, which are useful as immunogens, can be

CC obtained by introducing these DNAs into eucaryotic cells.

SQ Sequence 1086 BP; 223 A; 335 C; 358 G; 170 T;

alignment\_scores:  
 Quality: 49.00 Length: 10  
 Ratio: 4.900 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-08-653-294-26 x Q01822 ..

Align seg 1/1 to: Q01822 from: 1 to: 1086

11 ArgGluAsnLeuArgIleAlaLeuArgTyr 20

|||||  
 294 CGAGAGAACCTGCGGATCGCGCTCGCTAC 323

seq\_name: N\_Geneseq\_36:Q05693

seq\_documentation\_block:

ID Q05693 standard; DNA; 1089 BP.

AC Q05693;

DT 03-JAN-1991 (first entry)

DE HLA-B51 gene for production of monoclonal antibodies.

KW Allotype specific monoclonal anti-HLA antibodies; hybridomas;

KW transgenic animals; HLA-B51 gene; ss.

FH Key Location/Qualifiers

FT exon 1..73

FT /\*tag= a

FT /number=1

FT 74..343

FT exon

FT /\*tag= b  
 FT /number=2  
 FT /note="alpha 1-domain"  
 FT 344..619  
 FT /\*tag= c  
 FT /number=3  
 FT /note="alpha 2-domain"  
 FT 620..895  
 FT /\*tag= d  
 FT /number=4  
 FT /note="alpha 3-domain"  
 FT 896..1012  
 FT /\*tag= e  
 FT /number=5  
 FT 1013..1042  
 FT /\*tag= f  
 FT /number=6  
 FT 1043..1089  
 FT /\*tag= g  
 FT /number=7

EP-383183-A.

22-AUG-1990.

PD 07-FEB-1990; 102424.

PF 08-FEB-1989; JP-029313.

PR (OLYU ) OLYMPUS OPTICAL KK.

PA Takiguchi M;

PI WPI; 90-255479/34.

DR Allotype specific monoclonal anti- HLA antibodies prodn. - using

PT hybridomas derived from transgenic animals carrying HLA gene and

PT immunised with HLA antigen of different allotype

PS Disclosure; Fig 1 A-G; 20pp; English.

CC The human HLA-B51 gene was injected into fertilised mouse eggs and

CC then these introduced into the uterus of a pseudo pregnant mouse.

CC The young were tested to ensure incorporation of the gene into the

CC chromosome, and one of them mated 3 times with a normal male to

CC produce 16 young, seven of which carried the HLA-B51 gene.

CC The transgenic offspring were immunised with HLA antigen.

CC The spleen lymphocytes were fused with myeloma cells. Hybridomas

CC producing antibodies were selected.

CC See also Q05701.

SQ Sequence 1089 BP; 224 A; 335 C; 357 G; 173 T;

alignment\_scores:

Quality: 49.00 Length: 10

Ratio: 4.900 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-08-653-294-26 x Q05693 ..

Align seg 1/1 to: Q05693 from: 1 to: 1089

11 ArgGluAsnLeuArgIleAlaLeuArgTyr 20

|||||  
 295 CGAGAGAACCTGCGGATCGCGCTCGCTAC 324

seq\_name: N\_Geneseq\_36:Q05701

seq\_documentation\_block:

ID Q05701 standard; DNA; 1089 BP.

AC Q05701;

DT 03-JAN-1991 (first entry)

DE HLA-Bw52 gene for production of monoclonal antibodies.

KW Allotype specific monoclonal anti-HLA antibodies; hybridomas;

KW transgenic animals; HLA-Bw52 gene; ss.

FH Key Location/Qualifiers

FT exon 1..73

FT /\*tag= a

FT /number=1

FT 74..343

FT /\*tag= b

FT /number=2

FT

FT exon /note="alpha 1-domain"  
 344..619  
 FT /tag= c  
 FT /number=3  
 FT /note="alpha 2-domain"  
 620..895  
 FT /tag= d  
 FT /number=4  
 FT /note="alpha 3-domain"  
 896..1012  
 FT /tag= e  
 FT /number=5  
 FT exon 1013..1042  
 FT /tag= f  
 FT /number=6  
 FT exon 1043..1089  
 FT /tag= g  
 FT /number=7  
 PN EP-383183-A.  
 PD 22-AUG-1990.  
 PF 07-FEB-1990; 102424.  
 PR 08-FEB-1989; JP-029313.  
 PA (OLYU ) OLYMPUS OPTICAL KK.  
 PI Takiguchi M;  
 DR WPI; 90-255479/34.  
 PT Allotype specific monoclonal anti- HLA antibodies prodn. - using  
 PT hybridomas derived from transgenic animals carrying HLA gene and  
 PT immunised with HLA antigen of different allotype  
 PS Disclosure: Fig 1 A-G; 20pp; English.  
 CC The human HLA-Bw52 gene was introduced into mouse L cells and  
 CC then these cells used to immunise one of the transgenic mice  
 CC (See Q05693).  
 CC The spleen lymphocytes were fused with myeloma cells (P3x63-Ag8.653).  
 CC Hybridomas producing antibodies were selected.  
 SQ Sequence 1089 BP; 223 A; 336 C; 359 G; 171 T;

alignment\_scores:  
 Quality: 49.00 Length: 10  
 Ratio: 4.900 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000  
 alignment\_block:  
 US-08-653-294-26 x Q05701 ..  
 Align seg 1/1 to: Q05701 from: 1 to: 1089

11 ArgGluAsnLeuArgIleAlaLeuArgTyr 20  
 |||||  
 295 CGAGAGAACTGCGGATCGGCTCGCTAC 324

seq\_name: N\_Geneseq\_36:Q12114

seq\_documentation\_block:  
 ID Q12114 standard; DNA: 1089 BP.  
 AC Q12114;  
 DT 29-AUG-1991 (first entry)  
 DE HLA-Bw53 exon.  
 KW Human leukocyte antigen; probe; major histocompatibility complex;  
 KW MHC; class I; ss.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT cds 1..1089  
 FT /tag= a  
 PN J03112487-A.  
 PD 14-MAY-1991.  
 PF 22-SEP-1989; 247697.  
 PR 22-SEP-1989; JP-247697.  
 PA (OLYU ) OLYMPUS OPTICAL KK.  
 DR WPI; 91-182991/25.  
 DR P-PSDB; R12463.  
 PT HLA-Bw53 gene, DNA probe and transformant cells - used for  
 PT immunisation, identifying specificity of antiserum etc.

PS Claim 1; Page 1; lpp; Japanese.  
 CC Probes comprising part of the sequence can be used to identify  
 CC Class I genes. The DNA can be expressed for immunisation of  
 CC animals and prodn. of monoclonal antibodies specific for the  
 CC HLA-Bw53 antigen. See also J03112485 and J03112486. 174 T;  
 SQ Sequence 1089 BP; 222 A; 337 C; 356 G;

alignment\_scores:  
 Quality: 49.00 Length: 10  
 Ratio: 4.900 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-08-653-294-26 x Q12114 ..

Align seg 1/1 to: Q12114 from: 1 to: 1089

11 ArgGluAsnLeuArgIleAlaLeuArgTyr 20  
 |||||  
 295 CGAGAGAACTGCGGATCGGCTCGCTAC 324

seq\_name: N\_Geneseq\_36:T42608

seq\_documentation\_block:

ID T42608 standard; DNA: 285 BP.

AC T42608;

DT 07-AUG-1997 (first entry)

DE Synthetic gene, lqhiv encoding scorpion neurotoxin.

KW neurotoxin; pest control; insect; acarid; nematode; lqhiv; lqhvi;

KW venom; Leirus quinquistriatus hebraeus; scorpion; lepidptera;

KW baculovirus; arthropod; ds.

OS Synthetic.

FH Key Location/Qualifiers

FT cds 70..287

FT /tag= a

FT /note= "no start codon given"

PN W09636221-A2.

PD 21-NOV-1996.

PF 30-APR-1996; U06076.

PR 08-MAY-1995; US-435040.

PA (REGC ) UNIV CALIFORNIA.

PI Hammock BD, Herrmann R, Moskowitz H;

DR WPI; 97-011744/01.

DR P-PSDB; W06339.

PT Control of insects, acarids and nematodes - using novel scorpion

PT toxins or a combination of 2 or more insect toxins

PS Claim 1; Fig 1; 35pp; English.

CC A novel method for controlling pests selected from insects, acarids, and

CC nematodes, comprises treating the pests or their loci with at least two

CC different insect toxins, the source of the toxins being at least one

CC recombinant microbe, the toxins having non-overlapping binding sites at

CC an insect cellular membrane channel. The present sequence is a synthetic

CC gene for lqhiv toxin, derived from the venom of Leirus quinquistriatus

CC hebraeus (Scorpion). lqhiv is the most potent lepidpteros toxin

CC isolated from scorpion venom to date, whilst lqhvi (W06340) has weak

CC mammal toxicity. In particular the insect toxin genes are expressed

CC by baculovirus vectors, as the baculoviruses infect only arthropods,

CC therefore pose little or no risk to humans, plants or the environment.

SQ Sequence 285 BP; 67 A; 92 C; 70 G; 56 T;

alignment\_scores:  
 Quality: 47.00 Length: 19  
 Ratio: 2.938 Gaps: 0  
 Percent Similarity: 84.211 Percent Identity: 52.632

alignment\_block:

US-08-653-294-26 x T42608 ..

Align seg 1/1 to: T42608 from: 1 to: 285

1 TyrArgLeuAlaIleArgLeuAsnGluArgGluAsnLeuArgIleAl 17





PA (MEIP ) MEIJI MILK PROD CO LTD.  
PI Ito Y, Sasaki T, Sasaki Y;  
DR WPI: 92-316178/38.  
PT Plasmid pBULI derived from Lactobacillus delbrueckii strain - is  
PT lactic acid bacteria culture promotion vector and lactic acid  
PT bacteria-Escherichia coli shuttle vector  
PS Claim 4; Page 3; 34pp; Japanese.  
CC The sequence is a 1344 bp SmaI cleavage fragment of pBULI isolated  
CC from Lactobacillus delbrueckii subsp. bulgaricus M-878 (FERM  
CC BP-3758), a new strain of lactobacillus. The plasmid or its derivs.  
CC bacteria, e.g. in the prodn. of yoghurt. It may also be used as a  
CC lactic acid bacteria/E. coli shuttle vector.  
SQ Sequence 1344 BP; 392 A; 295 C; 356 G; 301 T;

alignment\_scores:  
Quality: 44.00 Length: 14  
Ratio: 3.385 Gaps: 0  
Percent Similarity: 92.857 Percent Identity: 57.143

alignment\_block:  
US-08-653-294-26 x Q28344 ..

Align seg 1/1 to: Q28344 from: 1 to: 1344

2 ArgLeuAlaIleArgLeuAsnGluArgGluAsnLeuArg 15  
||||| :|||:|||||:|||||:|||||:|||||:|||||  
148 GCCTTAGAATTACGATTGACGGACAGGAAGAGAAATTCCTG 189

seq\_name: N\_Geneseq\_36:T07447

seq\_documentation\_block:  
ID T07447 standard; DNA; 362 BP.  
AC T07447;  
DT 27-APR-1996 (first entry)  
DE B. subtilis aprE target DNA.  
KW DNA; nucleic acid; analyte; lateral flow detection; PCR;  
KW polymerase chain reaction; capture probe; hybridisation; aprE gene;  
KW ss.  
OS Bacillus subtilis.  
PN W09527081-AI.  
PD 12-OCT-1995.  
PF 30-MAR-1995; U03773.  
PR 31-MAR-1994; US-221769.  
PA (DUPO ) DU PONT DE NEMOURS & CO E I.  
PI Ebersole RC, Fitzpatrick-McElligott S, Hendrickson ER;  
PI Majarian WR, Payne MS, Rafalski JA;  
DR WPI: 95-358650/46.  
PT Detecting the presence of a nucleic acid analyte - by contact with a  
PT test strip having immobilised capture moieties.  
PS Example 1; Page 43; 82pp; English.  
CC A specific segment (T07447) of the B. subtilis aprE gene was amplified  
CC by PCR using primers 3 (T07448) and 4 (T07449) in the presence of  
CC biotin-14-dATP and digoxigenin-11-dUTP. Detection of the PCR product  
CC was performed using a lateral flow nucleic acid detection system in  
CC which anti-digoxigenin coupled to alkaline phosphatase was  
CC deposited on the 'application zone' of a nitrocellulose membrane,  
CC and avidin on the 'capture zone'. Amplified aprE target DNA was  
CC applied to the application zone and allowed to wick to the capture  
CC zone, where it was detected by colorimetry. In another method,  
CC target aprE DNA amplified in the presence of biotin-labeled dNTP  
CC was hybridised in solution with aprE DNA amplified in the presence  
CC of digoxigenin-labeled dNTP to demonstrate lateral flow detection of a  
CC bifunctional nucleic acid hybrid.  
SQ Sequence 362 BP; 99 A; 89 C; 82 G; 92 T;

alignment\_scores:  
Quality: 43.00 Length: 17  
Ratio: 2.867 Gaps: 0  
Percent Similarity: 88.235 Percent Identity: 52.941

alignment\_block:  
US-08-653-294-26 x T07447/rev ..

Align seg 1/1 to reverse of: T07447 from: 1 to: 362

3 LeuAlaIleArgLeuAsnGluArgGluAsnLeuArgIleAlaLeuArg 19  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
111 GTAGCCTTGAGAGTGAAGACGGCGCTTATTATTGAGAAATGCCATAAG 62  
19 g 19  
61 G 61

seq\_name: N\_Geneseq\_36:N80745

seq\_documentation\_block:

ID N80745 standard; DNA; 1220 BP.

AC N80745;  
DT 14-SEP-1990 (first entry)  
DE Coding region of the aprA gene encoding subtilisin in Bacillus subtilis  
DE Bacillus subtilis strain Q8127 (trpC2 leuA8 sacU200); subtilisin;  
KW thermal stability; pH stability; specific activity;  
KW substrate specificity; detergent; ss.  
OS Bacillus subtilis strain Q8127 (trpC2 leuA8 sacU200).  
FH Key Location/Qualifiers  
FT mat\_peptide 1..1143 /\*tag= a  
FT WO8808033-A.  
PN 20-OCT-1988.  
PD 28-MAR-1988; U01038.  
PF 10-APR-1988; US-036872.  
PR (AMGE-) Amgen Inc.  
PA Zurovski MM, Stabinsky Y, Levitt M;  
PI WPI: 88-307568/43.  
DR P-PSDB; P80744.  
PT New subtilisin analogues -  
PT have aminoacid(s) present in calcium binding site replaced by  
PT negatively charged aminoacid(s)  
PS Example 2; Table 1, Pages 22-24; 60pp; English.  
CC Genomic DNA was isolated from cells of B. subtilis strain Q8127 (trpC2  
CC leuA8 sacU200). The coding region of the aprA gene was sequenced and  
CC the results of the sequence are given in n80745. The specific identity  
CC of the initial 5 codons of the leader region is attributable to the  
CC report of Stahl, et al., J. Bacteriol., 158, 411-418, (1984) and Wong,  
CC et al P.N.A.S., 81, 1184-1188 (1984). There exist codon sequence  
CC differences from Stahl, et al., at amino acid positions 84 and 85.  
CC Specifically, Stahl, et al., reports a codon GTT (coding for valine) at  
CC amino acid position 84 while the codon GTA (also coding for valine)  
CC appears in n80745. Stahl, et al., also reports a codon AGC (coding for  
CC serine) at amino acid position 85 as opposed to the codon GCG (coding for  
CC alanine) in n80745. The patent concerns a novel subtilisin analogue  
CC which has an amino acid sequence of a naturally occurring Bacillus  
CC subtilisin which has been modified by having: one or more of the amino  
CC acids present in a calcium binding site of the naturally occurring  
CC Bacillus subtilisin replaced by a negatively charged amino acid, and  
CC one or more of any Asn-Gly sequence of the naturally occurring Bacillus  
CC subtilisin deleted or replaced by a different amino acid. Pref. it is  
CC an analogue of subtilisin Carlsberg, subtilisin DY, subtilisin BPN', an  
CC aprA subtilisin of B. subtilis or subtilisin from B. mesentericus. The  
CC subtilisin analogues exhibit improved thermal and pH stability,  
CC increased specific activity and broad substrate specificity thereby  
CC increasing the detergent of detergent formulations contg. such  
CC analogues.  
SQ Sequence 1220 BP; 355 A; 281 C; 283 G; 301 T;

alignment\_scores:  
Quality: 43.00 Length: 17  
Ratio: 2.867 Gaps: 0  
Percent Similarity: 88.235 Percent Identity: 52.941

alignment\_block:  
US-08-653-294-26 x N80745/rev ..

Align seg 1/1 to reverse of: N80745 from: 1 to: 1220

```

3 LeuAlaIleArgLeuAsnGluArgGluAsnLeuArgIleAlaLeuAr 19
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
381 GTAGCCTTGAGAGTGAAGCGCGCTTTAATTGAGAAATGCCATAAG 332
19 g 19
331 G 331

```

seq\_name: N\_Geneseq\_36:Q06587

```

seq_documentation_block:
ID_ Q06587 standard; DNA; 1220 BP.
AC Q06587;
DT 18-FEB-1991 (first entry)
DE aprA gene encoding subtilisin.
KW Detergents; depilatory tanning; serology; ds.
OS Bacillus subtilis.
FH Key Location/Qualifiers
FT conflict 568..573
FT /label= a
FT /note= "Stahl et al. J.Bacteriol.,158,411-418 (1984)"
FT cds 1..1143
FT /tag= b
FT mat_peptide 319..1143
FT /tag= c
PN EP-398539-A.
PD 22-NOV-1990.
PF 01-MAY-1990; 304715.
PR 17-MAY-1989; US-353124.
PA (AMGE-) AMGEN INC.
PI Zukowski MM, Narhi LO, Levitt M;
DR WPI: 90-350298/47.
DR P-PSDB; R07970.
PT Bacillus subtilin analogues - with improved pH thermal and oxidn.
PS Claim 20; Table 1; 39pp; English.
CC Modified analogues of subtilisin are useful in cleaning fabrics, and
CC have an improved resistance to oxidation, heat and pH extremes.
CC Analogues have one or more negative AAs present in the calcium
CC binding site (CBS) and may also be used in manufacture of protein
CC hydrolysates, and detection of incomplete Abs in serology.
SQ Sequence 1220 BP; 355 A; 283 C; 283 G; 301 T;

```

alignment\_scores:  
Quality: 43.00 Length: 17  
Ratio: 2.867 Gaps: 0  
Percent Similarity: 88.235 Percent Identity: 52.941

alignment\_block:

US-08-653-294-26 x Q06587/rev ..

Align seg 1/1 to reverse of: Q06587 from: 1 to: 1220

```

3 LeuAlaIleArgLeuAsnGluArgGluAsnLeuArgIleAlaLeuAr 19
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
381 GTAGCCTTGAGAGTGAAGCGCGCTTTAATTGAGAAATGCCATAAG 332
19 g 19
331 G 331

```

seq\_name: N\_Geneseq\_36:N70890

```

seq_documentation_block:
ID_ N70890 standard; DNA; 1220 BP.
AC N70890;
DT 26-FEB-1991 (first entry)
DE Thermostable Bacillus subtilisin encoded by aprA gene.

```

```

KW aprA gene; subtilisin; thermostable enzyme; protease; surfactant; ds.
OS Bacillus subtilis.
FH Key Location/Qualifiers
FT cds 1..1143
FT /tag= a
FT mat_peptide 319..1143
FT /tag= b
FT /label= aprA gene
PN W08704461-A.
PD 30-JUL-1987.
PF 07-JAN-1987; U00027.
PR 15-JAN-1986; US-819241.
PA (AMGE-) AMGEN.
PI Stabinsky Y, Zukowski M;
DR WPI: 87-221262/31.
DR P-PSDB; P70555.
PT Thermally stable and pH stable subtilisin analogues - produced by
PT deleting or replacing at least one of the asparagine-glycine
PT sequences.
PS Disclosure: Table 1; 56pp; English.
CC Modified subtilisin may be used in detergent compositions, it has
CC an improved pH and heat stability. See also P70556.
SQ Sequence 1220 BP; 355 A; 283 C; 281 G; 301 T;

```

alignment\_scores:  
Quality: 43.00 Length: 17  
Ratio: 2.867 Gaps: 0  
Percent Similarity: 88.235 Percent Identity: 52.941

alignment\_block:

US-08-653-294-26 x N70890/rev ..

Align seg 1/1 to reverse of: N70890 from: 1 to: 1220

```

3 LeuAlaIleArgLeuAsnGluArgGluAsnLeuArgIleAlaLeuAr 19
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
381 GTAGCCTTGAGAGTGAAGCGCGCTTTAATTGAGAAATGCCATAAG 332
19 g 19
331 G 331

```

seq\_name: N\_Geneseq\_36:Q03536

seq\_documentation\_block:

ID\_ Q03536 standard; DNA; 1499 BP.

```

AC Q03536;
DT 16-AUG-1990 (first entry)
DE Subtilisin gene.
KW Carboxyl hydrolase; subtilisin; neutral protease; ds.
OS Bacillus subtilis.
FH Key Location/Qualifiers
FT cds 137..1280
FT /tag= a
PN EP-357157-A.
PD 7-MAR-1990.
PF 22-JUN-1984; 202584.
PR 24-JUN-1983; US-507419.
PR 29-MAY-1984; US-614612.
PR 29-MAY-1984; US-614615.
PR 29-MAY-1984; US-614491.
PR 29-MAY-1984; US-614616.
PR 29-MAY-1984; US-614617.
PA (GETH) Genentech Inc.
PI Bott RR, Estell DA, Ferrari E, Henner DJ, Wells JA;
DR WPI: 90-068909/10.
DR P-PSDB; R03737.

```

```

PT Mutant prokaryotic carbonyl hydrolase enzymes -
PT obtd. by site-directed oligo-nucleotide mutagenesis, used in
PT food processing and cleaning industries.
PS Claim 16; Fig 7; 39pp; English.
CC Probe derived from subtilisin gene was used to isolate carbonyl

```

```

alignment_block:
US-08-653-294-26 x Q03536/rev ..
Align seg 1/1 to reverse of: Q03536 from: 1 to: 1499

3 LeuAlaIleArgLeuAsnGluArgAGluAsnLeuArgIleAlaLeuAr 19
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
517 GTAGCCTTGAGAGTGARGAGCCGCGCTTTAATTGGAATGCCATAAG 468

19 g 19
      |
467 G 467

```

**THIS PAGE BLANK (USPTO)**



Rhabditia; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.

REFERENCE 1 (bases 1 to 360)  
 AUTHORS Kohara.Y., Mitsuki.H., Nishigaki.A., Motohashi.T., Sugimoto.A. and Tabara.H.  
 TITLE Toward an expression map of the C.elegans genome  
 JOURNAL Unpublished (1994)  
 COMMENT Contact: Yuji Kohara  
 Gene Library Lab  
 National Institute of Genetics  
 Yata 1111, Mishima, Shizuoka 411, Japan  
 Tel: 81-559-81-6854  
 Fax: 81-559-81-6855  
 Email: ykohara@lab.nig.ac.jp.

FEATURES  
 source  
 1..360  
 /organism="Caenorhabditis elegans"  
 /strain="CB1489 him-8(e1489)"  
 /db\_xref="taxon:6239"  
 /clone="yk196"  
 /clone\_lib="Yuji Kohara unpublished cDNA"  
 /sex="hermaphrodite, male"  
 /tissue\_type="whole animal"  
 /dev\_stage="varied"  
 BASE COUNT 116 a 85 c 79 g 80 t  
 ORIGIN

alignment\_scores  
 Quality: 55.00 Length: 17  
 Ratio: 3.667 Gaps: 0  
 Percent Similarity: 88.235 Percent Identity: 70.588

alignment\_block  
 US-08-653-294-26 x D35321 ..  
 Align seg 1/1 to: D35321 from: 1 to: 360  
 2 ArgLeuAlaIleArgLeuAsnGluArgGluAsnLeuArgIleAlaLe 18  
 ||| |||:::|||||:::|||||:::|||||:::|||||::: ||  
 79 AGATCCCGGTTAGGTGACGGAAGAAGATCCCAATCTGAGGATTATCT 128

18 u 18  
 129 G 129

seq\_name: gb\_est1:D35579

seq\_documentation\_block: 360 bp mRNA EST 08-AUG-1994  
 LOCUS D35579  
 DEFINITION CELK022G9F Yuji Kohara unpublished cDNA Caenorhabditis elegans cDNA  
 clone yk22g9 5', mRNA sequence.

ACCESSION D35579  
 VERSION D35579.1 GI:527030  
 KEYWORDS EST.  
 SOURCE Caenorhabditis elegans.  
 ORGANISM Caenorhabditis elegans.

REFERENCE 1 (bases 1 to 360)  
 AUTHORS Kohara.Y., Mitsuki.H., Nishigaki.A., Motohashi.T., Sugimoto.A. and Tabara.H.  
 TITLE Toward an expression map of the C.elegans genome  
 JOURNAL Unpublished (1994)  
 COMMENT Contact: Yuji Kohara  
 Gene Library Lab  
 National Institute of Genetics  
 Yata 1111, Mishima, Shizuoka 411, Japan  
 Tel: 81-559-81-6854  
 Fax: 81-559-81-6855  
 Email: ykohara@lab.nig.ac.jp.

FEATURES  
 source  
 1..360  
 /organism="Caenorhabditis elegans"

/strain="CB1489 him-8(e1489)"  
 /db\_xref="taxon:6239"  
 /clone="yk22g9"  
 /clone\_lib="Yuji Kohara unpublished cDNA"  
 /sex="hermaphrodite, male"  
 /tissue\_type="whole animal"  
 /dev\_stage="varied"  
 BASE COUNT 106 a 84 c 87 g 77 t 6 others  
 ORIGIN

alignment\_scores  
 Quality: 55.00 Length: 17  
 Ratio: 3.667 Gaps: 0  
 Percent Similarity: 88.235 Percent Identity: 70.588

alignment\_block  
 US-08-653-294-26 x D35579 ..  
 Align seg 1/1 to: D35579 from: 1 to: 360  
 2 ArgLeuAlaIleArgLeuAsnGluArgGluAsnLeuArgIleAlaLe 18  
 ||| |||:::|||||:::|||||:::|||||:::|||||::: ||  
 24 AGATCCCGGTTAGGTGACGGAAGAAGATCCCAATCTGAGGATTATCT 73

18 u 18  
 74 G 74

seq\_name: gb\_est17:C69731

seq\_documentation\_block: 360 bp mRNA EST 23-SEP-1997  
 LOCUS C69731  
 DEFINITION C69731 Yuji Kohara unpublished cDNA Caenorhabditis elegans cDNA  
 clone yk368h10 5', mRNA sequence.

ACCESSION C69731  
 VERSION C69731.1 GI:2431087  
 KEYWORDS EST.  
 SOURCE Caenorhabditis elegans.  
 ORGANISM Caenorhabditis elegans.

REFERENCE 1 (bases 1 to 360)  
 AUTHORS Kohara.Y., Motohashi.T., Tabara.H., Watanabe.H., Sugimoto.A., Sano.M., Miyata.A. and Nishigaki.A.  
 TITLE Expression map of the C.elegans genome  
 JOURNAL Unpublished (1996)  
 COMMENT On Sep 12, 1996 this sequence version replaced gi:1316439.  
 Contact: Yuji Kohara  
 Gene Library Lab  
 National Institute of Genetics  
 Yata 1111, Mishima, Shizuoka 411, Japan  
 Tel: 81-559-81-6854  
 Fax: 81-559-81-6855  
 Email: ykohara@lab.nig.ac.jp.

FEATURES  
 source  
 1..360  
 /organism="Caenorhabditis elegans"  
 /strain="CB1489 him-8(e1489)"  
 /db\_xref="taxon:6239"  
 /clone="yk368h10"  
 /clone\_lib="Yuji Kohara unpublished cDNA"  
 /sex="hermaphrodite, male"  
 /tissue\_type="whole animal"  
 /dev\_stage="varied"

BASE COUNT 109 a 90 c 85 g 76 t  
 ORIGIN

alignment\_scores  
 Quality: 55.00 Length: 17  
 Ratio: 3.667 Gaps: 0  
 Percent Similarity: 88.235 Percent Identity: 70.588

seq\_name: gb\_est17:C69958

```

/clone_lib="C:\IOWAG\DNA1"
/lab_host="E. coli XL2 Blue MRF"
/notes="Vector: pBluescript II (SK-); Site:1: EcorV; C. parvum (IOWA isolate) genomic DNA was hydrodynamically sheared to produce fragments having a tight size distribution between 2-4 kb by Dr. Yvonne Thorstenson of the Stanford DNA Sequencing and Technology Center (http://sequence-www.stanford.edu/group/techdev/shear.htm). The randomly sheared gDNA was chromatographed on Sephacryl S-400 to remove any small fragments and DNA eluting in the void volume was subcloned into an EcoR V-digested, alkaline phosphatase-treated pBluescript II (SK-) vector and transformed into E. coli strain XL2 Blue MRF". Recombinant clones from the first plating of the library were selected for sequence analysis using T3 and T7 primers."
BASE COUNT      252 a 130 c 108 g 300 t 1 others
ORIGIN

alignment_scores:
    Quality: 51.00      Length: 20
    Ratio: 3.400      Gaps: 0
    Percent Similarity: 75.000      Percent Identity: 55.000

alignment_block:
US-08-653-294-26 x AQ855602/rev ..

Align seg 1/1 to reverse of: AQ855602 from: 1 to: 791

1 TyArgLeuAlaIleArgLeuAsnGluArgGluAsnLeuArgIleAl 17
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
674 TATAGGAATTCATCAGGGAARACCGAGGTCGATAATATGAGGATTC 625

17 aLeuArgTyr 20
|||||:::
624 ACTCAGATTT 615

seq_name: gb_est8:c03945

seq_documentation_block:
LOCUS      C03945
DEFINITION C03945 Human heart cDNA (Ynakamura) Homo sapiens cDNA clone
3NHC2454, mRNA sequence.
ACCESSION C03945
VERSION    C03945.1 GI:1467196
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
            Tanaka,T., Ogiwara,A., Uchiyama,I., Takagi,T., Yazaki,Y. and
            Nakamura,Y.
            Construction of a normalized directionally cloned cDNA library fr
            adult heart and analysis of 3040 clones by partial sequencing
            Genomics 35 (1), 231-235 (1996)
96299762

On Oct 24, 1995 this sequence version replaced gi:1040105.
Contact: Yusuke Nakamura
Institute of Medical Science
University of Tokyo
4-6-1, Shirokanedai, Minato-ku, Tokyo 108, Japan
Tel: 81-3-5449-5372
Fax: 81-3-5449-5433
Email: yusuke@ims.u-tokyo.ac.jp.
Location/Qualifiers
1..232
/organism="Homo sapiens"
/db_xref="caxon:9606"
/clone="3NHC2454"
/clone_lib="Human heart cDNA (Ynakamura)"

FEATURES
Source

```



```
/dev_stage="adult"  
/note="organ: heart; normalized directionally cloned cdna  
from adult heart"
```

BASE COUNT 55 a 77 c 68 g 32 t  
ORIGIN

alignment\_scores:  
Quality: 49.00 Length: 10  
Ratio: 4.900 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-08-653-294-26 x C03945 ..

Align seg 1/1 to: C03945 from: 1 to: 232

```
11 ArgGluAsnLeuArgIleAlaLeuArgTyr 20  
|||||  
40 CGAGAGAACCTGCGGATCGCGCTCCGCTAC 69
```

seq\_name: gb\_est10:AA151891

seq\_documentation\_block: 255 bp mRNA EST 10-DEC-1996  
LOCUS AA151891  
DEFINITION zoolf06.r1 Stratagene colon (#937204) Homo sapiens cdna clone  
IMAGE:566435 5' similar to gb:M15497.cdsl HLA CLASS I  
HISTOCOMPATIBILITY ANTIGEN, A-24(A-9) A\*2401 (HUMAN);, mRNA  
sequence.

ACCESSION AA151891  
VERSION AA151891.1 GI:1720754  
KEYWORDS EST.  
SOURCE human.

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
AUTHORS

Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B.,  
Chissoe, S., Dietrich, N., DuBuque, T., Favellio, A., Gish, W.,  
Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N.,  
Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L.,  
Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J.,  
Trevisan, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R.,  
and Marra, M.

TITLE Generation and analysis of 280,000 human expressed sequence tags  
Genome Res. 6 (9), 807-828 (1996)

JOURNAL 9704478  
MEDLINE  
COMMENT On May 8, 1995 this sequence version replaced gi:800234.

Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.

Trace considered overall poor quality  
Seq primer: -28M13 rev2 from Amersham

High quality sequence stop: 1.

FEATURES  
source

1. .255  
/organism="Homo sapiens"  
/db\_xref="GDB:4590888"  
/db\_xref="taxon:9606"  
/clone="IMAGE:566435"

/clone\_lib="Stratagene colon (#937204)"  
/lab\_host="SOUR cells (kanamycin resistant)"

/note="organ: colon; Vector: pBluescript SK-; Site: 1:  
EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:  
Oligo dT. T-84 colonic epithelial cell line. Average  
insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor  
sequence: 5' GAATTCGGCAGAG 3' -3' adaptor sequence: 5'  
CTCGAGTATTTTTTTTTTTT 3'"

BASE COUNT 57 a 70 c 75 g 44 t 9 others  
ORIGIN

alignment\_scores:  
Quality: 49.00 Length: 10  
Ratio: 4.900 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-08-653-294-26 x AA151891 ..

Align seg 1/1 to: AA151891 from: 1 to: 255

```
11 ArgGluAsnLeuArgIleAlaLeuArgTyr 20  
|||||  
77 CGAGAGAACCTGCGGATCGCNCCTCCGCTAC 106
```

seq\_name: gb\_est11:AA263158

seq\_documentation\_block: 283 bp mRNA EST 02-JUL-1998

LOCUS AA263158  
DEFINITION PM0534 KGI-a Lambda zap Express cdna library Homo sapiens CDNA 5',  
mRNA sequence.

ACCESSION AA263158

VERSION AA263158.1 GI:1898964

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

AUTHORS

Claudio, J.O., Liew, C.C., Dempsey, A.A., Cukerman, E., Stewart, A.K.,  
Na, E., Atkins, H.I., Iscoe, N.N. and Hawley, R.G.

TITLE Identification of sequence-tagged transcripts differentially  
expressed within the human hematopoietic hierarchy

JOURNAL Genomics 50 (1), 44-52 (1998)

MEDLINE 98292493

COMMENT On May 5, 1995 this sequence version replaced gi:797810.

Contact: Hawley RG

Oncology Research Laboratories

The Toronto Hospital

CRC-424, 67 College St., Toronto, Ontario M5G 2M1, Canada

Tel: 416 3403834

Fax: 416 3403453

Email: r.hawley@utoronto.ca

Similar to M58636 MHC class I HLA-B\* gene. Clone was randomly

picked from KGIa primary library.

Seq primer: 5' GAAATTAACCTCACTAAAGG 3'

High quality sequence stop: 283.

FEATURES

source

1. .283  
Location/Qualifiers

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone\_lib="KGI-a Lambda zap Express cdna library"

/cell\_type="promyeloblast"

/cell\_line="KGI-a"

/note="vector: Lambda zap Express (Stratagene); Site: 1:  
EcoRI; Site 2: XhoI; Unidirectional cloning sites:

EcoRI-XhoI. mRNA was purified from KGI-a cell line, cDNA

was synthesized using an XhoI-OligodT linker primer. EcoRI

adaptors were ligated, followed by digestion with XhoI for

directional cloning into predigested Lambda zap Express"

64 a 91 c 88 g 40 t

BASE COUNT  
ORIGIN

alignment\_scores:

Quality: 49.00 Length: 10

Ratio: 4.900 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

```

US-08-653-294-26 x AA263158 ..
Align seg 1/1 to: AA263158 from: 1 to: 283

11 ArgGluAsnLeuArgIleAlaLeuArgTyr 20
|||||
120 CGAGAGAACCTCGGATCGCGCTCGCTAC 149

seq_name: gb_est6:D82221

seq_documentation_block:
LOCUS D82221 375 bp mRNA EST 09-FEB-1996
DEFINITION HUMHBC4626 Human pancreatic islet Homo sapiens cDNA similar to
HLA-B mRNA sequence.
ACCESSION D82221
VERSION D82221.1 GI:1183739
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 375)
AUTHORS Takeda,J.
TITLE Human pancreatic islet ESTs
JOURNAL Unpublished (1995)
COMMENT On Apr 14, 1993 this sequence version replaced gi:785255.
Contact: Jun Takeda
Institute for Molecular and Cellular Regulation, Gunma University
3-39-15 Showa-machi, Maebashi Gunma 371, Japan
Tel: 272-20-8856
Fax: 272-20-8896
Email: jtakeda@sb.gunma-u.ac.jp.
Location/Qualifiers
1. 375
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Human pancreatic islet"
/notes="Vector: Lambda ZAPII; Site.1: Eco RI; Site.2: Xho
I; mRNA was prepared from normal adult human islets. cDNA
was directionally synthesized from the Xho I in the vector
to the EcoRI site. cDNA was size fractionated to remove
sequences <1000 bp in size."
BASE COUNT 75 a 124 c 118 g 55 t 3 others
ORIGIN

alignment_scores:
Quality: 49.00 Length: 10
Ratio: 4.900 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-26 x D82221 ..
Align seg 1/1 to: D82221 from: 1 to: 375

11 ArgGluAsnLeuArgIleAlaLeuArgTyr 20
|||||
306 CGAGAGAACCTCGGATCGCGCTCGCTAC 335

seq_name: gb_est26:AI359260

seq_documentation_block:
LOCUS AI359260 618 bp mRNA EST 15-FEB-1999
DEFINITION qy27b07.x1 NCI_CGAP_Brn23 Homo sapiens cDNA clone IMAGE:2013205 3'
similar to gb:D32129 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
AW-66(A-10) A*6601 ALPHA (HUMAN);, mRNA sequence.
ACCESSION AI359260
VERSION AI359260.1 GI:4110881
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute / National Institute of Neurological
Disorders and Stroke, Brain Tumor Genome Anatomy Project
(CGAP/BTGP), Tumor Gene Index
Unpublished (1998)
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,
Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html

Insert Length: 691 Std Error: 0.00
Seq primer: -40UP from Gbco
High quality sequence stop: 458.
Location/Qualifiers
1. 618
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="IMAGE:2013205"
/clone_lib="NCI_CGAP_Brn23"
/tissue_type="glioblastoma (pooled)"
/lab_host="DH10B"
/notes="Organ: brain; Vector: pT7T3D-Pac (Pharmacia) with a
modified polylinker; Site.1: Not I; Site.2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5',
TGTTCACCATCTGAAGTCGAGCGCGCATATCTTTTTTTTTTTTTTTTTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT7T3 vector.
Library is normalized, and was constructed by Bento
Soares and M.Fatima Bonaldo."
BASE COUNT 128 a 171 c 182 g 137 t
ORIGIN

alignment_scores:
Quality: 49.00 Length: 10
Ratio: 4.900 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-26 x AI359260/rev ..
Align seg 1/1 to reverse of: AI359260 from: 1 to: 618

11 ArgGluAsnLeuArgIleAlaLeuArgTyr 20
|||||
319 CGAGAGAACCTCGGATCGCGCTCGCTAC 290

seq_name: gb_gss15:AQ649574

seq_documentation_block:
LOCUS AQ649574 698 bp DNA GSS 22-JUN-1999
DEFINITION Sheared DNA-28D9.TF Sheared DNA Trypanosoma brucei genomic clone
Sheared DNA-28D9, genomic survey sequence.
ACCESSION AQ649574
VERSION AQ649574.1 GI:5142760
KEYWORDS GSS.
SOURCE Trypanosoma brucei.
ORGANISM Trypanosoma brucei
Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;

REFERENCE 1 (bases 1 to 698)
AUTHORS El-Sayed,N., Zhao,S., Zhao,H., Gill,S., Suh,E., Malek,J., Fujii,C.,

```

Gerrard,C., Leech,V., de Jong,P., Ullu,E., Melville,S.,  
 Doneison,J., Fraser,C. and Adams,M.  
**TITLE** Determination of clone end sequences from Trypanosoma brucei GUTAT  
**JOURNAL** 10.1 sheared DNA library  
**COMMENT** Unpublished (1999)  
 On Mar 23, 1999 this sequence version replaced gi:3325266.  
**Other\_GSSs:** Sheared DNA-28D9.TR  
**Contact:** Najib M. El-Sayed  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 0200  
 Fax: 301 838 0208  
 Email: nelsayed@tigr.org  
 Clones are derived from the Trypanosoma brucei GUTAT 10.1 sheared  
 DNA library constructed at TIGR. Clones will be available for  
 distribution through ATCC. Sheared DNA end sequences search page:  
<http://www.tigr.org/tdb/mdb/cbdb/>.  
**Seq primer:** M13-Forward  
**Class:** Shotgun.

**FEATURES** Location/Qualifiers  
 source 1..698

/organism="Trypanosoma brucei"  
 /strain="TREU927/4 Gutat 10.1"  
 /db\_xref="taxon:5691"  
 /clone="Sheared DNA-28D9"  
 /clone\_lib="Sheared DNA"

/notes="Vector: pUC19; Site:1: SmaI; Constructed at The  
 Institute for Genomic Research (TIGR), Rockville, MD.  
 Genomic DNA isolated from a cloned population of  
 Trypanosoma brucei (TREU927/4 Gutat 10.1) was mechanically  
 sheared to give a tight size distribution (approx 2 kb).  
 The v + i method used for the library construction is  
 described in detail in Smith, H.O. and Venter J.C.  
 (Making small insert libraries for whole genome shotgun  
 sequencing projects. in Genome Sequencing: A Practical  
 Approach, eds. M. Vaudin and B. Borell, Oxford University  
 Press, 1999)."

**BASE COUNT** 199 a 159 c 223 g 117 t  
**ORIGIN**

**alignment\_scores:**  
 Quality: 49.00 Length: 19  
 Ratio: 3.062 Gaps: 0  
**Percent Similarity:** 84.211 **Percent Identity:** 52.632

**alignment\_block:**

US-08-653-294-26 x AQ649574/rev ..

Align seg 1/1 to reverse of: AQ649574 from: 1 to: 698

2 ArgLeuAlaIleArgLeuAsnGluArgArgGluAsnLeuArgIleAlaLe 18  
 :::::::::::::::::::::::::::: ::::::::::: ||  
 389 CAGCTCTGTGTCGGCTCAATGAGAAGTGACTCCGCTTCGCTTCGCT 340

18 uArgTyr 20

|||||

339 CCGCTAC 333

seq\_name: gb\_est31:AI696864

seq\_documentation\_block:

LOCUS AI696864 748 bp mRNA EST 03-JUN-1999  
**DEFINITION** wc74h11.x1 NCI\_CGAP\_Pan1 Homo sapiens cDNA clone IMAGE:2324421 3'  
 similar to gb:M28205 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,  
 B\*51(B\*5) B\*5101 ALPHA (HUMAN);, mRNA sequence.

**ACCESSION** AI696864

**VERSION** AI696864.1 GI:4984764

**KEYWORDS** EST.

**SOURCE** human.

**ORGANISM** Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria: Primates; Catarrhini; Hominoidea; Homo.  
**REFERENCE** 1 (bases 1 to 748)  
**AUTHORS** NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
**TITLE** National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
**JOURNAL** Unpublished (1997)  
**COMMENT** On Mar 16, 1998 this sequence version replaced gi:2961758.  
**Contact:** Robert Strausberg, Ph.D.  
 Tel: (301) 496-1550  
 Email: Robert.Strausberg@nih.gov  
 Life Technologies catalog #: 11548-013  
 DNA sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
[www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)

**Seq primer:** -40UP from Gibco  
**High quality sequence stop:** 424.

**FEATURES** Location/Qualifiers

source

1..748  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:2324421"  
 /clone\_lib="NCI\_CGAP\_Pan1"  
 /tissue\_type="adenocarcinoma"  
 /lab\_host="DH10B"  
 /notes="Organ: pancreas; Vector: pCMV-SPORT6; Site:1: SalI;  
 Site:2: NotI; Cloned unidirectionally. Primer: Oligo dt.  
 Average insert size 1.72 kb. Life Technologies catalog #:  
 11548-013"

**BASE COUNT** 169 a 227 c 237 g 108 t 7 others  
**ORIGIN**

**alignment\_scores:**  
 Quality: 49.00 Length: 10  
 Ratio: 4.900 Gaps: 0  
**Percent Similarity:** 100.000 **Percent Identity:** 100.000

**alignment\_block:**

US-08-653-294-26 x AI696864 ..

Align seg 1/1 to: AI696864 from: 1 to: 748

11 ArgGluAsnLeuArgIleAlaLeuArgTyr 20

|||||

137 CGAGAGACCTGCGGATCGCGCTCGCTAC 166

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2000, 04:05:43 ; Search time 133.56 seconds  
(without alignments)  
3.547 Million cell updates/sec

Title: US-08-653-294-27

Perfect score: 102

Sequence: 1 YGRNLSRRESLRNLRY 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 188963 seqs, 23686106 residues

Total number of hits satisfying chosen parameters: 188963

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : A\_Geneseq\_36.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	102	100.0	20	1 R92913	HLA-B7 CTL modul
2	102	100.0	20	1 R95415	HLA-B7.84-75-84 Pa
3	102	100.0	20	1 W33790	Peptide B7.84-75/7
4	102	100.0	20	1 W33797	Peptide B7.84-75/7
5	51	50.0	10	1 R41209	Peptide fragment o
6	51	50.0	10	1 R83061	HLA-B7 CTL modul
7	51	50.0	10	1 W07515	T-cell modulating
8	51	50.0	10	1 W33786	Peptide B7.75-84 t
9	51	50.0	10	1 W33796	Peptide B7.75-84 t
10	51	50.0	25	1 R41207	Peptide fragment o
11	51	50.0	25	1 R83073	HLA-Bw62 CTL modul
12	51	50.0	25	1 R95431	HLA-B7.60-84. Comp
13	51	50.0	25	1 R95419	HLA-Bw62.60-84. Co
14	51	50.0	362	1 R12464	HLA-B35 antigen. H
15	47.5	46.6	605	1 W62838	Glycine max antimi
16	46	45.1	509	1 W26792	Mouse TIE-2 recept
17	44	43.1	20	1 R92910	HLA-B2702 CTL modu
18	44	43.1	20	1 W33793	Peptide B2702.84-7
19	44	43.1	25	1 R41206	Peptide fragment o
20	44	43.1	25	1 R83072	HLA-Bw46 CTL modul
21	44	43.1	25	1 R95418	HLA-Bw46.60-84. Co
22	44	43.1	366	1 R12465	HLA-C exon Cb-1. H
23	44	43.1	366	1 R12466	HLA-C exon Cb-2. H
24	44	43.1	366	1 Y07033	Breast cancer asso
25	43	42.2	430	1 R20642	Glyphosate oxidore
26	43	42.2	431	1 R22262	Glyphosate oxidore
27	42	41.2	20	1 R92909	HLA-B2702 CTL modu
28	42	41.2	20	1 R92908	HLA-B2702 CTL modu
29	42	41.2	20	1 W33791	Peptide B2702.84-7
30	42	41.2	20	1 W33792	Peptide B2702.84-7
31	42	41.2	2233	1 W48711	HP1V-3 JS isolate
32	42	41.2	2233	1 W48712	HP1V-3 FRh1 cp45 v
33	42	41.2	2233	1 W48713	HP1V-3 vero cp45 v
34	41	40.2	217	1 W46424	Human macrophage s

35 41 40.2 219 1 W98326 H. pylori GHPO 144  
36 41 40.2 274 1 P80911 Consensus sequence  
37 41 40.2 705 1 R66597 Human L5/3 tumour  
38 41 40.2 705 1 R66598 Human L5/3 tumour  
39 41 40.2 705 1 W14266 Human L5/3 partial  
40 41 40.2 705 1 W14267 Human L5/3 partial  
41 41 40.2 710 1 W07691 Macrophage stimula  
42 41 40.2 711 1 R66602 Human L5/3 tumour  
43 41 40.2 711 1 W07692 Macrophage stimula  
44 41 40.2 711 1 W14270 Human growth facto  
45 41 40.2 711 1 W82789 Human MSP protein.

#### ALIGNMENTS

RESULT 1

R92913  
ID R92913 standard; peptide; 20 AA.  
AC R92913;  
DT 16-MAY-1996 (first entry)  
DE HLA-B7 CTL modulating peptide (B7.84-75/75-84)  
KW Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;  
KW immunosuppressant; graft versus host disorder; transplantation; therapy;  
KW class I MHC; HLA-B7.  
OS Synthetic.  
PN W09526979-A1.  
PD 12-OCT-1995.  
PF 05-APR-1995; U04349.  
PR 05-APR-1994; US-222851.  
PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
PI Clayberger C, Krensky AM, Parham P;  
DR WPI: 95-358582/46.  
PT Extension of acceptance period of transplants from MHC unmatched  
PT donor hosts - using Class I B75-84 MHC antigen of the recipient  
PT host  
PS Example 15; Page 36; 80pp; English.  
CC R83061-R83085, R83090-R83096 and R92907-R92913 represent fragments of  
CC class I major histocompatibility complex (MHC) antigens. This sequence  
CC is an inverted dimer of residues 75-84 of the alpha-1 domain of the class  
CC I MHC HLA-B7. These sequences can be used to extend the period of  
CC acceptance by a recipient of a transplant from an MHC unmatched donor.  
CC The peptides are administered to a patient in conjunction with a  
CC subtherapeutic amount of an immunosuppressant. This is administered to  
CC the patient for a limited period of time (compared to the lifetime  
CC administration for current treatments). The peptides particularly  
CC modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)  
CC of the patient.  
SQ Sequence 20 AA;

Query Match 100.0%; Score 102; DB 1; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1.3e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YGRNLSRRESLRNLRY 20  
| | | | | | | | | | | | | | | | | | | | | |  
DB 1 YGRNLSRRESLRNLRY 20

RESULT 2

R95415  
ID R95415 standard; peptide; 20 AA.  
AC R95415;  
DT 12-NOV-1996 (first entry)  
DE HLA-B7.84-75-84 Palindrome.  
KW HLA; p74; alpha1-helix; human-leucocyte-associated antigen; inhibitor;  
KW T-cell lysate; membrane protein; mammal; heat shock protein; Hsc70; APC;  
KW B cell; calcium influx; cytotoxic T lymphocyte; CTL; differentiation;  
KW cytolysis; antigen presenting cell.  
OS Synthetic.  
PN W09513288-A1.  
PD 18-MAY-1995.

PF 10-NOV-1994; U12985.  
 PR 10-NOV-1993; US-150493.  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI Clayberger C, Krensky AM;  
 DR WPI: 95-194027/25.  
 PT Compens. comprising lymphoid surface membrane proteins - which may  
 PT inhibit cytolytic activity and differentiation of CTLs.  
 PS Example; Page 18; 29pp; English.  
 CC R95413, and R95415-R95431 represent palindromes and fragments of  
 CC human-leucocyte-associated antigens. This sequence represents the  
 CC HLA-B\*75-84 palindromic. These sequences can be used to isolate  
 CC the protein p74 from a T-cell lysate. p74 is a T-cell surface membrane  
 CC protein associated with T-cell activation in mammalian T-cells, and is  
 CC also immunologically cross reactive with the heat shock protein Hsc70.  
 CC p74 is found in a limited number of cell types, but is particularly  
 CC expressed on B and T cells. p74 can be isolated by lysis of a suitable  
 CC cell with an amphoteric detergent, and then passed through an affinity  
 CC column containing a covalently bound HLA-B2702 palindromic peptide.  
 CC Compositions comprising the extracellular fragment of p74 combined with  
 CC HLA-B2702.60-84 (see R95416), induces calcium influx, and inhibits  
 CC cytotoxic T lymphocyte (CTL) differentiation or cytotoxicity. Candidate  
 CC compounds can be screened for their effect on the cytolytic activity of  
 CC T-cells, by combining them with the extracellular portion of p74 and  
 CC determining the amount of binding between the candidate compound and p74.  
 CC Modulation of CTL activity can be inhibited in a cellular composition  
 CC containing T-cells and antigen presenting cells (APCs), by adding to the  
 CC mix the extracellular portion of p74, in an amount sufficient to compete  
 CC with p74 for the binding of the p74 ligand.  
 CC Sequence 20 AA;  
 SQ

Query Match 100.0%; Score 102; DB 1; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-09;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YGRNLSRRESRLNRYG 20  
 |||||  
 DB 1 YGRNLSRRESRLNRYG 20

## RESULT 3

ID W33790 standard; peptide; 20 AA.  
 AC W33790;  
 DT 19-JUN-1998 (first entry)  
 DE Peptide B7.84-75/75-84 tested for immunomodulating activity.  
 KW Immunomodulating dimer; immunosuppressant drug; CTL activation;  
 KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;  
 KW rejection.  
 OS Synthetic.  
 OS Homo sapiens.  
 PN WO9744351-A1.  
 PD 27-NOV-1997.  
 PF 22-MAY-1997; U08689.  
 PR 24-MAY-1996; US-653294.  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI Beulow R, Clayberger C, Krensky AM;  
 DR WPI: 98-086530/08.  
 PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B  
 PT alpha-1 domain, used for preventing rejection of transplants or  
 PT treating autoimmune diseases  
 PS Example 1; Page 19; 41pp; English.  
 CC Peptides W33784-98 and W33778-9 were assayed for their immunomodulating  
 CC activity. A peptide-type compound or variant is claimed which has  
 CC immunomodulating activity, including the N-terminal acylated and/or  
 CC C-terminal amidated or esterified forms of up to 60 amino acids, where  
 CC the peptide-type compound comprises the formula: A-B, where A, B =  
 CC (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or V; aa77 =  
 CC D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a hydrophobic or  
 CC small amino acid; aa82 = R or L; aa83 = G or R; and aa represents amino  
 CC acid. The sequence in the brackets may optionally be absent or truncated  
 CC at any peptide type bond within the brackets. The compounds comprise  
 CC amino acid sequences related to a Class I HLA-B alpha-1 domain (positions  
 CC 79-84). They can be used to inhibit cytotoxic T-lymphocytes (CTL) from  
 CC undesirably attacking cells in a host or in vitro. They can also be  
 CC used in combination with antigenic peptides or proteins of interest to  
 CC activate CTLs. They can also inhibit the proliferation of T cells in  
 CC response to anti-CD3. The peptide can be used for preventing rejection  
 CC of transplants or for treating autoimmune diseases, e.g. diabetes,  
 CC rheumatoid arthritis and lupus erythematosus. The products can also be  
 CC used for detection and diagnosis.  
 CC Sequence 20 AA;  
 SQ

Query Match 100.0%; Score 102; DB 1; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-09;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YGRNLSRRESRLNRYG 20  
 |||||  
 DB 1 YGRNLSRRESRLNRYG 20

CC 79-84). They can be used to inhibit cytotoxic T-lymphocytes (CTL) from  
 CC undesirably attacking cells in a host or in vitro. They can also be  
 CC used in combination with antigenic peptides or proteins of interest to  
 CC activate CTLs. They can also inhibit the proliferation of T cells in  
 CC response to anti-CD3. The peptide can be used for preventing rejection  
 CC of transplants or for treating autoimmune diseases, e.g. diabetes,  
 CC rheumatoid arthritis and lupus erythematosus. The products can also be  
 CC used for detection and diagnosis.  
 CC Sequence 20 AA;  
 SQ

Query Match 100.0%; Score 102; DB 1; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-09;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YGRNLSRRESRLNRYG 20  
 |||||  
 DB 1 YGRNLSRRESRLNRYG 20

## RESULT 4

ID W33797 standard; peptide; 20 AA.  
 AC W33797;  
 DT 19-JUN-1998 (first entry)  
 DE Peptide B7.84-75/75-84 tested for immunomodulating activity.  
 KW Immunomodulating dimer; immunosuppressant drug; CTL activation;  
 KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;  
 KW rejection.  
 OS Synthetic.  
 OS Homo sapiens.  
 PN WO9744351-A1.  
 PD 27-NOV-1997.  
 PF 22-MAY-1997; U08689.  
 PR 24-MAY-1996; US-653294.  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI Beulow R, Clayberger C, Krensky AM;  
 DR WPI: 98-086530/08.  
 PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B  
 PT alpha-1 domain, used for preventing rejection of transplants or  
 PT treating autoimmune diseases  
 PS Example 1; Page 19; 41pp; English.  
 CC Peptides W33784-98 and W33778-9 were assayed for their immunomodulating  
 CC activity. A peptide-type compound or variant is claimed which has  
 CC immunomodulating activity, including the N-terminal acylated and/or  
 CC C-terminal amidated or esterified forms of up to 60 amino acids, where  
 CC the peptide-type compound comprises the formula: A-B, where A, B =  
 CC (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or V; aa77 =  
 CC D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a hydrophobic or  
 CC small amino acid; aa82 = R or L; aa83 = G or R; and aa represents amino  
 CC acid. The sequence in the brackets may optionally be absent or truncated  
 CC at any peptide type bond within the brackets. The compounds comprise  
 CC amino acid sequences related to a Class I HLA-B alpha-1 domain (positions  
 CC 79-84). They can be used to inhibit cytotoxic T-lymphocytes (CTL) from  
 CC undesirably attacking cells in a host or in vitro. They can also be  
 CC used in combination with antigenic peptides or proteins of interest to  
 CC activate CTLs. They can also inhibit the proliferation of T cells in  
 CC response to anti-CD3. The peptide can be used for preventing rejection  
 CC of transplants or for treating autoimmune diseases, e.g. diabetes,  
 CC rheumatoid arthritis and lupus erythematosus. The products can also be  
 CC used for detection and diagnosis.  
 CC Sequence 20 AA;  
 SQ

```

RESULT 5
R41209
ID R41209 standard; peptide; 10 AA.
AC R41209; 50.0%; Score 51; DB 1; Length 10;
DE Peptide fragment of Class I HLA peptide.
KW Human leukocyte antigen; HLA; peptide; transplantation; neoplasia;
OS parasitic disease; cytotoxic T lymphocyte; modulation.
PN WO9317699-A.
PD 16-SEP-1993.
PF 25-FEB-1993; U01758.
PR 02-MAR-1992; US-844716.
PA (STRD ) UNIV LELAND STANFORD JUNIOR.
PI Clayberger CA, Krensky AM;
DR WPI; 93-303134/38.
PT New peptides(s) based on Class I HLA antigen domains - used for
PT modulating cytotoxic T-lymphocyte activity towards targets
PS Claim 11; Page 54; 61pp; English.
CC The peptide is used to modulate cytotoxic T-lymphocyte (CTL)
CC activity, either by inhibition of stimulation. It can be used
CC for inhibiting CTL toxicity in transplantations, for inducing CTL
CC activity in parasitic diseases and neoplasia and in studies on viral
CC infection. The peptide can also be used for identifying CTLs which
CC bind to it and removing subsets of CTLs from a T-cell composition.
CC This peptide sequence is more commonly found within larger peptide
CC compounds of not more than 30 amino acids in length.
SQ Sequence 10 AA;

Query Match 50.0%; Score 51; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.043;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RESLNRLRGY 20
DB 1 RESLNRLRGY 10

RESULT 6
R83061
ID R83061 standard; peptide; 10 AA.
AC R83061;
DE HLA-B7 CTL modulating peptide (B7.75-84).
KW Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;
KW immunosuppressant; graft versus host disorder; transplantation; therapy;
OS Synthetic.
PN WO9526979-A1.
PD 12-OCT-1995.
PF 05-APR-1995; U04349.
PR 03-APR-1994; US-222851.
PA (STRD ) UNIV LELAND STANFORD JUNIOR.
PI Clayberger C, Krensky AM, Parham P;
DR WPI; 95-358582/46.
PT Extension of acceptance period of transplants from MHC unmatched
PT donor hosts - using Class I B75-84 MHC antigen of the recipient
PT host
PS Claim 13; Page 66; 80pp; English.
CC This sequence represents a fragment of a class I major histocompatibility
CC complex (MHC) antigen. This sequence corresponds to residues 75-84 of
CC the alpha-1 domain of the class I MHC HLA-B7. This sequence, and the
CC peptide fragments represented by R83062-R83085, R83090-R83096 and
CC R83097-R83098 can be used to extend the period of acceptance by a
CC recipient of a transplant from an MHC unmatched donor. The peptides are
CC administered to a patient in conjunction with a subtherapeutic amount of
CC an immunosuppressant. This is administered to the patient for a limited
CC period of time (compared to the lifetime administration for current
CC treatments). The peptides particularly modulate (or inhibit) the
CC activity of the cytotoxic T lymphocytes (CTLs) of the patient.
SQ Sequence 10 AA;

Query Match 50.0%; Score 51; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.043;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RESLNRLRGY 20
DB 1 RESLNRLRGY 10

RESULT 7
W07515
ID W07515 standard; peptide; 10 AA.
AC W07515;
DE 04-AUG-1997 (first entry)
DE T-cell modulating peptide #4.
KW mammal; major histocompatibility complex; MHC class I; antigen; perforin;
KW insulin-dependent diabetes mellitus; multiple sclerosis; inflammation;
KW rheumatoid arthritis; psoriasis; pemphigus vulgaris; Sjogren's disease;
KW thyroid disease; Hashimoto's thyroiditis; myasthenia gravis; granzyme;
KW autologous target cell; cytokine release; T cell activation; therapy.
OS Synthetic.
PN WO9635443-A1.
PD 14-NOV-1996.
PF 05-APR-1996; U04710.
PR 12-MAY-1995; US-440504.
PA (SANG-) SANGSTAT MEDICAL CORP.
PI Buelow R;
DR WPI; 96-518410/51.
PT Treatment of auto-immune disease by admin. of peptide(s) corresp. to
PT major histocompatibility complex antigens - esp. for delaying onset
PT of clinical symptoms of insulin dependent diabetes by modulating T
PT cell mediated attack on target cells
PS Claim 7; Page 20; 24pp; English.
CC W07512-W07518 represent T-cell modulating peptides that can be used in
CC the method of the invention. These sequences are based on a portion of
CC the generic peptide corresponding to residues 70-91 of the alpha-1 domain
CC of the major histocompatibility complex (MHC) class I antigen (see
CC W07510). The method is for affecting the course of an autoimmune disease
CC involving T-cell mediated destruction of tissue in mammals. These
CC peptides are used especially to treat insulin-dependent diabetes
CC mellitus, preferably being administered during the pre-clinical stage to
CC delay onset of the disease. Other diseases that can be treated are
CC multiple sclerosis, rheumatoid arthritis, psoriasis, pemphigus vulgaris,
CC Sjogren's disease, thyroid disease, Hashimoto's thyroiditis, myasthenia
CC gravis, etc. The peptides modulate T-cell mediated attack on autologous
CC target cells, and may also reduce inflammation, swelling, and release of
CC cytokines, perforins, granzymes etc. associated with T cell activation.
SQ Sequence 10 AA;

Query Match 50.0%; Score 51; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.043;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RESLNRLRGY 20
DB 1 RESLNRLRGY 10

RESULT 8
W33786
ID W33786 standard; peptide; 10 AA.
AC W33786;
DE 19-JUN-1998 (first entry)
DE Peptide B7.75-84 tested for immunomodulating activity.
KW immunomodulating dimer; immunosuppressant drug; CTL activation;
KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;
KW rejection.
OS Synthetic.
PN Homo sapiens.
PN WO9744351-A1.

Query Match 50.0%; Score 51; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.043;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RESLNRLRGY 20
DB 1 RESLNRLRGY 10

```

PD 27-NOV-1997.  
 PF 22-MAY-1997; U08689.  
 PR 24-MAY-1996; US-653294.  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI Beulow R, Clayberger C, Krensky AM;  
 DR WPI: 98-086530/08.  
 PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B  
 PT alpha-1 domain, used for preventing rejection of transplants or  
 PT treating autoimmune diseases  
 PS Example 1: Page 19; 41pp; English.  
 CC Peptides W33784-98 and W33778-9 were assayed for their immunomodulating  
 CC activity. A peptide-type compound or variant is claimed which has  
 CC immunomodulating activity, including the N-terminal acylated and/or  
 CC C-terminal amidated or esterified forms of up to 60 amino acids, where  
 CC the peptide-type compound comprises the formula: A-B, where A, B =  
 CC (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or V; aa77 =  
 CC D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a hydrophobic or  
 CC small amino acid; aa82 = R or L; aa83 = G or R; and aa represents amino  
 CC acid. The sequence in the brackets may optionally be absent or truncated  
 CC at any peptide type bond within the brackets. The compounds comprise  
 CC amino acid sequences related to a Class I HLA-B alpha domain (positions  
 CC 79-84). They can be used to inhibit cytotoxic T-lymphocytes (CTL) from  
 CC undesirably attacking cells in a host or in vitro. They can also be  
 CC used in combination with antigenic peptides or proteins of interest to  
 CC activate CTLs. They can also inhibit the proliferation of T cells in  
 CC response to anti-CD3. The peptide can be used for preventing rejection  
 CC of transplants or for treating autoimmune diseases, e.g. diabetes,  
 CC rheumatoid arthritis and lupus erythematosus. The products can also be  
 CC used for detection and diagnosis.  
 SQ Sequence 10 AA;

Query Match 50.0%; Score 51; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.043;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RESLRNLRCY 20  
 | | | | | | | | | |  
 Db 1 RESLRNLRCY 10

RESULT 9  
 W33796  
 ID W33796 standard; peptide: 10 AA.  
 AC W33796;  
 DT 19-JUN-1998 (first entry)  
 DE Peptide B7.75-84 tested for immunomodulating activity.  
 KW Immunomodulating dimer; immunosuppressant drug; CTL activation;  
 KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;  
 KW rejection.  
 OS Synthetic.  
 OS Homo sapiens.  
 PN WO9744351-A1.  
 PD 27-NOV-1997.  
 PF 22-MAY-1997; U08689.  
 PR 24-MAY-1996; US-653294.  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI Beulow R, Clayberger C, Krensky AM;  
 DR WPI: 98-086530/08.  
 PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B  
 PT alpha-1 domain, used for preventing rejection of transplants or  
 PT treating autoimmune diseases  
 PS Example 1: Page 19; 41pp; English.  
 CC Peptides W33784-98 and W33778-9 were assayed for their immunomodulating  
 CC activity. A peptide-type compound or variant is claimed which has  
 CC immunomodulating activity, including the N-terminal acylated and/or  
 CC C-terminal amidated or esterified forms of up to 60 amino acids, where  
 CC the peptide-type compound comprises the formula: A-B, where A, B =  
 CC (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or V; aa77 =  
 CC D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a hydrophobic or  
 CC small amino acid; aa82 = R or L; aa83 = G or R; and aa represents amino  
 CC acid. The sequence in the brackets may optionally be absent or truncated  
 CC at any peptide type bond within the brackets. The compounds comprise

CC amino acid sequences related to a Class I HLA-B alpha domain (positions  
 CC 79-84). They can be used to inhibit cytotoxic T-lymphocytes (CTL) from  
 CC undesirably attacking cells in a host or in vitro. They can also be  
 CC used in combination with antigenic peptides or proteins of interest to  
 CC activate CTLs. They can also inhibit the proliferation of T cells in  
 CC response to anti-CD3. The peptide can be used for preventing rejection  
 CC of transplants or for treating autoimmune diseases, e.g. diabetes,  
 CC rheumatoid arthritis and lupus erythematosus. The products can also be  
 CC used for detection and diagnosis.  
 SQ Sequence 10 AA;

Query Match 50.0%; Score 51; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.043;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RESLRNLRCY 20  
 | | | | | | | | | |  
 Db 1 RESLRNLRCY 10

RESULT 10  
 R41207  
 ID R41207 standard; peptide: 25 AA.  
 AC R41207;  
 DT 15-MAR-1994 (first entry)  
 DE Peptide fragment of Class I HLA peptide.  
 KW Human leukocyte antigen; HLA; peptide; transplantation; neoplasia;  
 KW parasitic disease; cytotoxic T lymphocyte; modulation.  
 OS Synthetic.  
 PN WO9317699-A.  
 PD 16-SEP-1993.  
 PF 25-FEB-1993; U01758.  
 PR 02-MAR-1992; US-844716.  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI Clayberger CA, Krensky AM;  
 DR WPI: 93-303134/38.  
 PT New peptide(s) based on Class I HLA antigen domains - used for  
 PT modulating cytotoxic T-lymphocyte activity towards targets  
 PS Claim 10; Page 54; 61pp; English.  
 CC The peptide (or a fragment of at least 10 amino acids, joined at at  
 CC least one terminus to a sequence other than that of wild type HLA  
 CC antigen) is used to modulate cytotoxic T-lymphocyte (CTL) activity,  
 CC either by inhibition or stimulation. It can be used for  
 CC inhibiting CTL toxicity in transplantations, for inducing CTL  
 CC activity in parasitic diseases and neoplasia and in studies on viral  
 CC infection. The peptide can also be used for identifying CTLs which  
 CC bind to it and removing subsets of CTLs from a T-cell composition.  
 SQ Sequence 25 AA;

Query Match 50.0%; Score 51; DB 1; Length 25;  
 Best Local Similarity 100.0%; Pred. No. 0.11;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RESLRNLRCY 20  
 | | | | | | | | | |  
 Db 16 RESLRNLRCY 25

RESULT 11  
 R83073  
 ID R83073 standard; peptide: 25 AA.  
 AC R83073;  
 DT 16-MAR-1996 (first entry)  
 DE HLA-Bw62 CTL modulating peptide (Bw62.60-84).  
 KW Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;  
 KW immunosuppressant; graft versus host disorder; transplantation; therapy;  
 KW class I MHC; HLA-Bw62.  
 OS Synthetic.  
 PN WO9528979-A1.  
 PD 12-OCT-1995.  
 PF 05-APR-1995; U04349.



PR 05-APR-1994; US-222851.  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI Clayberger C, Krensky AM, Parham P;  
 WPI: 95-358582/46.  
 PT Extension of acceptance period of transplants from MHC unmatched  
 donor hosts - using Class I B\*9-84 MHC antigen of the recipient  
 PT host  
 PS Example 13; Page 32; 80pp; English.  
 CC R83061-R83085, R83090-R83096 and R92907-R92913 represent fragments of  
 CC class I major histocompatibility complex (MHC) antigens. This sequence  
 CC corresponds to residues 60-84 of the alpha-1 domain of the class I MHC  
 CC HLA-B\*62. These sequences can be used to extend the period of acceptance  
 CC by a recipient of a transplant from an MHC unmatched donor. The peptides  
 CC are administered to a patient in conjunction with a subtherapeutic amount  
 CC of an immunosuppressant. This is administered to the patient for a  
 CC limited period of time (compared to the lifetime administration for  
 CC current treatments). The peptides particularly modulate (or inhibit) the  
 CC activity of the cytotoxic T lymphocytes (CTLs) of the patient.  
 SQ Sequence 25 AA;

Query Match 50.0%; Score 51; DB 1; Length 25;  
 Best Local Similarity 100.0%; Pred. No. 0.11;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RESLRNLRGY 20  
 DB 16 RESLRNLRGY 25  
 |||||

RESULT 12  
 ID R95431 standard; peptide; 25 AA.  
 AC R95431;  
 DT 12-NOV-1996 (first entry)  
 DE HLA-B\*7.60-84.  
 KW HLA; p74; alpha-helix; human-leucocyte-associated antigen; inhibitor;  
 KW T-cell lysate; membrane protein; mammal; heat shock protein; Hsc70; APC;  
 KW B cell; calcium influx; cytotoxic T lymphocyte; CTL; differentiation;  
 KW cytolysis; antigen presenting cell.  
 OS Synthetic.  
 PN W09513288-A1.  
 PD 18-MAY-1995.  
 PF 10-NOV-1994; U12985.  
 PR 10-NOV-1993; US-150493.  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI Clayberger C, Krensky AM;  
 WPI: 95-194027/25.  
 DR Compens. comprising lymphoid surface membrane proteins - which may  
 PT inhibit cytolytic activity and differentiation of CTLs.  
 PS Example; Page 12; 29pp; English.  
 CC R95431, and R95415-R95431 represent palindromes and fragments of  
 CC human-leucocyte-associated antigens. This sequence represents the  
 CC HLA-B\*7.60-84. These sequences can be used to isolate the protein p74  
 CC from a T-cell lysate. p74 is a T-cell surface membrane protein  
 CC associated with T-cell activation in mammalian T-cells, and is also  
 CC immunologically cross reactive with the heat shock protein Hsc70. p74 is  
 CC found in a limited number of cell types, but is particularly expressed on  
 CC B and T cells. p74 can be isolated by lysis of a suitable cell with an  
 CC amphoteric detergent, and then passed through an affinity column  
 CC containing a covalently bound HLA-B2702 palindromic peptide.  
 CC Compositions comprising the extracellular fragment of p74 combined with  
 CC HLA-B2702.60-84 (see R95416), induces calcium influx, and inhibits  
 CC cytotoxic T lymphocyte (CTL) differentiation or cytolysis. Candidate  
 CC compounds can be screened for their effect on the cytolytic activity of  
 CC T-cells, by combining them with the extracellular portion of p74 and  
 CC determining the amount of binding between the candidate compound and p74.  
 CC Modulation of CTL activity can be inhibited in a cellular composition  
 CC containing T-cells and antigen presenting cells (APCs), by adding to the  
 CC mix the extracellular portion of p74, in an amount sufficient to compete  
 CC with p74 for the binding of the p74 ligand.  
 SQ Sequence 25 AA;

Query Match 50.0%; Score 51; DB 1; Length 25;  
 Best Local Similarity 100.0%; Pred. No. 0.11;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RESLRNLRGY 20  
 DB 16 RESLRNLRGY 25  
 |||||

RESULT 14  
 ID R12464 standard; Protein; 362 AA.  
 AC R12464;  
 DT 29-AUG-1991 (first entry)  
 DE HLA-B\*35 antigen.  
 KW Human leukocyte antigen; probe; major histocompatibility complex;  
 KW MHC; class I.  
 OS Homo sapiens.  
 PN J03112486-A.

Query Match 50.0%; Score 51; DB 1; Length 25;  
 Best Local Similarity 100.0%; Pred. No. 0.11;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RESLRNLRGY 20  
 DB 16 RESLRNLRGY 25  
 |||||

RESULT 13  
 ID R95419 standard; peptide; 25 AA.  
 AC R95419;  
 DT 12-NOV-1996 (first entry)  
 DE HLA-B\*62.60-84.  
 KW HLA; p74; alpha-helix; human-leucocyte-associated antigen; inhibitor;  
 KW T-cell lysate; membrane protein; mammal; heat shock protein; Hsc70; APC;  
 KW B cell; calcium influx; cytotoxic T lymphocyte; CTL; differentiation;  
 KW cytolysis; antigen presenting cell.  
 OS Synthetic.  
 PN W09513288-A1.  
 PD 18-MAY-1995.  
 PF 10-NOV-1994; U12985.  
 PR 10-NOV-1993; US-150493.  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI Clayberger C, Krensky AM;  
 WPI: 95-194027/25.  
 DR Compens. comprising lymphoid surface membrane proteins - which may  
 PT inhibit cytolytic activity and differentiation of CTLs.  
 PS Example; Page 9; 29pp; English.  
 CC R95413, and R95415-R95431 represent palindromes and fragments of  
 CC human-leucocyte-associated antigens. This sequence represents the  
 CC HLA-B\*62.60-84. These sequences can be used to isolate the protein p74  
 CC from a T-cell lysate. p74 is a T-cell surface membrane protein  
 CC associated with T-cell activation in mammalian T-cells, and is also  
 CC immunologically cross reactive with the heat shock protein Hsc70. p74 is  
 CC found in a limited number of cell types, but is particularly expressed on  
 CC B and T cells. p74 can be isolated by lysis of a suitable cell with an  
 CC amphoteric detergent, and then passed through an affinity column  
 CC containing a covalently bound HLA-B2702 palindromic peptide.  
 CC Compositions comprising the extracellular fragment of p74 combined with  
 CC HLA-B2702.60-84 (see R95416), induces calcium influx, and inhibits  
 CC cytotoxic T lymphocyte (CTL) differentiation or cytolysis. Candidate  
 CC compounds can be screened for their effect on the cytolytic activity of  
 CC T-cells, by combining them with the extracellular portion of p74 and  
 CC determining the amount of binding between the candidate compound and p74.  
 CC Modulation of CTL activity can be inhibited in a cellular composition  
 CC containing T-cells and antigen presenting cells (APCs), by adding to the  
 CC mix the extracellular portion of p74, in an amount sufficient to compete  
 CC with p74 for the binding of the p74 ligand.  
 SQ Sequence 25 AA;

Query Match 50.0%; Score 51; DB 1; Length 25;  
 Best Local Similarity 100.0%; Pred. No. 0.11;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RESLRNLRGY 20  
 DB 16 RESLRNLRGY 25  
 |||||

RESULT 14  
 ID R12464 standard; Protein; 362 AA.  
 AC R12464;  
 DT 29-AUG-1991 (first entry)  
 DE HLA-B\*35 antigen.  
 KW Human leukocyte antigen; probe; major histocompatibility complex;  
 KW MHC; class I.  
 OS Homo sapiens.  
 PN J03112486-A.

PD 14-MAY-1991.  
 PF 22-SEP-1989; 247697.  
 PR 22-SEP-1989; JP-247697.  
 PA (OLYU ) OLYMPUS OPTICAL KK.  
 DR WPI; 91-182991/25.  
 DR N-PSDB; Q12115.  
 PT HLA-B35 gene - used in DNA probe and transformant cells for  
 PT immunising animals, for developing monoclonal antibody.  
 PS Claim 1; Page 1; lpp; Japanese.  
 CC Probes comprising part of the sequence encoding this sequence can  
 CC be used to identify Class I genes. The DNA can be expressed for  
 CC immunisation of animals and prodn. of monoclonal antibodies  
 CC specific for the HLA-B35 antigen. See also J03112485 and J03112487.  
 SQ Sequence 362 AA;

Query Match 50.0%; Score 51; DB 1; Length 362;  
 Best Local Similarity 100.0%; Pred. No. 1.8;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RESLRNLRCY 20  
 DB 99 RESLRNLRCY 108

## RESULT 15

W62838  
 ID W62838 standard; Protein; 605 AA.  
 AC W62838;  
 DT 27-OCT-1998 (first entry)  
 DE Glycine max antimicrobial protein.  
 KW antimicrobial protein; infestation; control.  
 OS Glycine max.  
 PN W09827805-A1.  
 PD 02-JUL-1998.  
 PF 22-DEC-1997; AU0874.  
 PR 20-DEC-1996; AU-004275.  
 PA (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.  
 PI Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;  
 DR WPI; 98-377279/32.  
 PT Novel anti-microbial protein from e.g. Macadamia integrifolia -  
 PT useful for controlling microbial infestations of plants or mammals  
 PS Claim 1; Page 63-65; 96pp; English.  
 CC The sequence is that of an antimicrobial protein which can  
 CC be used to control microbial infestations in plants and mammalian  
 CC animals.  
 SQ Sequence 605 AA;

Query Match 46.6%; Score 47.5; DB 1; Length 605;  
 Best Local Similarity 47.8%; Pred. No. 11;  
 Matches 11; Conservative 3; Mismatches 6; Indels 3; Gaps 1;

QY 1 YGR---LNRLSERRESLRNLRCY 20  
 DB 211 YGRIRVLQRFNQSPQLQNLRDY 233

Search completed: February 8, 2000, 04:05:43  
 Job time: 9360 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 7, 2000, 18:04:38 ; Search time 111.22 Seconds  
(without alignments)  
8.482 Million cell updates/sec

Title: US-08-653-294-27

Perfect score: 102

Sequence: 1 YGRNRLSRERSLNLRGY 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 142080 seqs, 47169319 residues

Total number of hits satisfying chosen parameters: 142080

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : PIR\_62.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	51	50.0	137	2	I38875 MHC class I antige
2	51	50.0	137	2	I38876 MHC class I antige
3	51	50.0	137	2	I38860 MHC class I antige
4	51	50.0	137	2	I38874 MHC class I antige
5	51	50.0	181	2	I39188 MHC cell surface g
6	51	50.0	270	1	HLHU40 MHC class I histoc
7	51	50.0	274	2	I68774 MHC HLA-B39 chain
8	51	50.0	274	2	S24439 class I histocompa
9	51	50.0	300	2	I68701 cell surface antig
10	51	50.0	308	2	I36956 MHC class I lympho
11	51	50.0	350	2	I68747 MHC HLA B*1 - huma
12	51	50.0	350	2	I34308 class I histocompa
13	51	50.0	354	2	S24436 class I histocompa
14	51	50.0	354	2	S24437 class I histocompa
15	51	50.0	354	2	S24438 class I histocompa
16	51	50.0	354	2	S24440 class I histocompa
17	51	50.0	354	2	S24433 class I histocompa
18	51	50.0	358	2	S03538 MHC class I histoc
19	51	50.0	361	2	I54418 MHC class I histoc
20	51	50.0	362	1	HLHUB7 MHC class I histoc
21	51	50.0	362	2	S77966 MHC class I histoc
22	51	50.0	362	2	G01230 MHC class I histoc
23	51	50.0	362	2	S16789 MHC class I histoc
24	51	50.0	362	2	A45880 MHC class I histoc
25	51	50.0	362	2	I37519 MHC class I histoc
26	51	50.0	362	2	I37520 MHC class I histoc
27	51	50.0	362	2	I37522 MHC class I histoc
28	51	50.0	362	2	I84431 MHC HLA-B8 chain -
29	51	50.0	362	2	I72755 HLA-B*5602 - human
30	51	50.0	362	2	I84488 lymphocyte antigen

31	51	50.0	362	2	I72752 HLA-B*5501 - human
32	51	50.0	362	2	I72753 HLA-B*5502 - human
33	51	50.0	362	2	I72754 HLA-B*5601 - human
34	51	50.0	362	2	I38437 MHC class I histoc
35	51	50.0	362	2	I37492 HLA-B alpha-chain
36	51	50.0	362	2	I34505 lymphocyte antigen
37	51	50.0	362	2	I56130 HLA-B*5401 - human
38	51	50.0	362	2	I36962 MHC class I protei
39	51	50.0	362	2	I37475 HLA-B35 variant -
40	51	50.0	362	2	I56149 lymphocyte antigen
41	51	50.0	362	2	I59655 lymphocyte antigen
42	51	50.0	362	2	I59651 lymphocyte antigen
43	51	50.0	362	2	I61865 MHC HLA-Bw42, HLA-
44	51	50.0	362	2	I61859 MHC HLA-B14 chain
45	51	50.0	362	2	I61860 MHC HLA-B18 chain

ALIGNMENTS

RESULT 1

I38875

MHC class I antigen - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 07-Jun-1996 #sequence\_revision 07-Jun-1996 #text\_change 23-Jul-1999

C:Accession: I38875

R:Garber, T.L.; Butler, L.M.; Trachtenberg, E.A.; Erlich, H.A.; Rickards, O.; De Stef

Immunogenetics 42, 19-27, 1995

A:Title: HLA-B alleles of the Cayapa of Ecuador: new B39 and B15 alleles.

A:Reference number: I38860; MUID:95317819

A:Accession: I38875

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-137 <RES>

A:Cross-references: EMBL:U15639; NID:G930332; PIDN:AAA74046.1; PID:G930333

C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 50.0%; Score 51; DB 2; Length 137;  
Best Local Similarity 100.0%; Pred. No. 0.91;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RESLRNLRGY 20

Db 40 RESLRNLRGY 49

RESULT 2

I38876

MHC class I antigen - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 07-Jun-1996 #sequence\_revision 07-Jun-1996 #text\_change 23-Jul-1999

C:Accession: I38876

R:Garber, T.L.; Butler, L.M.; Trachtenberg, E.A.; Erlich, H.A.; Rickards, O.; De Stef

Immunogenetics 42, 19-27, 1995

A:Title: HLA-B alleles of the Cayapa of Ecuador: new B39 and B15 alleles.

A:Reference number: I38860; MUID:95317819

A:Accession: I38876

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-137 <RES>

A:Cross-references: EMBL:U15640; NID:G930334; PIDN:AAA74047.1; PID:G930335

C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 50.0%; Score 51; DB 2; Length 137;  
Best Local Similarity 100.0%; Pred. No. 0.91;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RESLRNLRGY 20

Db 40 RESLRNLRGY 49

```

RESULT 3
I38860
MHC class I antigen - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change 23-Jul-1999
C:Accession: I38860
R:Garber, T.L.; Butler, L.M.; Trachtenberg, E.A.; Erlich, H.A.; Rickards, O.; De Stefani
Immunogenetics 42, 19-27, 1995
A:Title: HLA-B alleles of the Cayapa of Ecuador: new B39 and B15 alleles.
A:Reference number: I38860; MUID:95317819
A:Accession: I38860
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-137 <RES>
A:Cross-references: EMBL:U14756; NID:g930328; PIDN:AA50171.1; PID:g930329
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 50.0%; Score 51; DB 2; Length 137;
Best Local Similarity 100.0%; Pred. No. 0.91; Indels 0; Gaps 0;
Matches 10; Conservative 0; Mismatches 0;

Qy 11 RESLNRLGY 20
|||||
Db 40 RESLNRLGY 49

RESULT 4
I38874
MHC class I antigen - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change 23-Jul-1999
C:Accession: I38874
R:Garber, T.L.; Butler, L.M.; Trachtenberg, E.A.; Erlich, H.A.; Rickards, O.; De Stefani
Immunogenetics 42, 19-27, 1995
A:Title: HLA-B alleles of the Cayapa of Ecuador: new B39 and B15 alleles.
A:Reference number: I38860; MUID:95317819
A:Accession: I38874
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-137 <RES>
A:Cross-references: EMBL:U15638; NID:g930330; PIDN:AAA74045.1; PID:g930331
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 50.0%; Score 51; DB 2; Length 137;
Best Local Similarity 100.0%; Pred. No. 0.91; Indels 0; Gaps 0;
Matches 10; Conservative 0; Mismatches 0;

Qy 11 RESLNRLGY 20
|||||
Db 40 RESLNRLGY 49

RESULT 5
I59188
MHC cell surface glycoprotein - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 23-Jul-1999
C:Accession: I59188
R:Bronson, S.K.; Pei, J.; Taillon-Miller, P.; Chorney, M.J.; Geraghty, D.E.; Chaplin, D.
Proc. Natl. Acad. Sci. U.S.A. 88, 1676-1680, 1991
A:Title: Isolation and characterization of yeast artificial chromosome clones linking th
A:Reference number: I59188; MUID:91156671
A:Accession: I59188
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-181 <RES>
A:Cross-references: GB:M59841; NID:g187697; PIDN:AAA59623.1; PID:g187698
C:Genetics:
A:Gene: GDB:HLA-B
A:Cross-references: GDB:120048; OMIM:142830

Query Match 50.0%; Score 51; DB 2; Length 137;
Best Local Similarity 100.0%; Pred. No. 0.91; Indels 0; Gaps 0;
Matches 10; Conservative 0; Mismatches 0;

Qy 11 RESLNRLGY 20
|||||
Db 40 RESLNRLGY 49

RESULT 6
HLH40
MHC class I histocompatibility antigen HLA-B40 alpha chain - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 28-Aug-1985 #sequence_revision 28-Aug-1985 #text_change 02-Sep-1997
C:Accession: A02186
R:Lopez de Castro, J.A.; Bragado, R.; Strong, D.M.; Strominger, J.L.
Biochemistry 22, 3961-3969, 1983
A:Title: Primary structure of papain-solubilized human histocompatibility antigen HLA
A:Reference number: A02186; MUID:84000412
A:Accession: A02186
A:Molecule type: protein
A:Residues: 1-270 <LOP>
C:Comment: Comparison of a number of class I allelic sequences (-B40 vs. -B7 and -A2
y restricted to the alpha-1 and alpha-2 domains. The most conspicuous clustering of v
ossible alloantigenic determinants of these antigens.
C:Genetics:
A:Gene: GDB:HLA-B
A:Cross-references: GDB:120048; OMIM:142830
A:Map position: 6p21.3-6p21.3
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology
C:Keywords: duplication; glycoprotein; heterodimer; membrane protein; transplantation
F:1-90/Domain: alpha-1 <EX1>
F:91-181/Domain: alpha-2 <EX2>
F:195-260/Domain: immunoglobulin homology <IMM>
F:86/Binding site: carbohydrate (Asn) (covalent) #status experimental
F:101-163,202-258/Disulfide bonds: #status predicted

Query Match 50.0%; Score 51; DB 1; Length 270;
Best Local Similarity 100.0%; Pred. No. 1.8; Indels 0; Gaps 0;
Matches 10; Conservative 0; Mismatches 0;

Qy 11 RESLNRLGY 20
|||||
Db 75 RESLNRLGY 84

RESULT 7
I68774
MHC HLA-B39 chain - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change 23-Jul-1999
C:Accession: I68774
R:Mueller, C.A.; Engler-Blum, G.; Gekeler, V.; Steiert, I.; Weiss, E.; Schmidt, H.
Immunogenetics 30, 200-207, 1989
A:Title: Genetic and serological heterogeneity of the supertypic HLA-B locus specific
A:Reference number: I54463; MUID:89379286
A:Accession: I68774
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-274 <RES>
A:Cross-references: GB:M29865; NID:g187676; PIDN:AAA36223.1; PID:g187677
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 50.0%; Score 51; DB 2; Length 274;
Best Local Similarity 100.0%; Pred. No. 1.8;

```

```

A:Map position: 6p21.3-6p21.3
A:Introns: 90/1
A:Superfamily: class I histocompatibility antigen; immunoglobulin homology
C:Keywords: glycoprotein

```

```

Query Match 50.0%; Score 51; DB 2; Length 181;
Best Local Similarity 100.0%; Pred. No. 1.2; Indels 0; Gaps 0;
Matches 10; Conservative 0; Mismatches 0;

```

```

Qy 11 RESLNRLGY 20
|||||
Db 74 RESLNRLGY 83

```

```

RESULT 6
HLH40
MHC class I histocompatibility antigen HLA-B40 alpha chain - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 28-Aug-1985 #sequence_revision 28-Aug-1985 #text_change 02-Sep-1997
C:Accession: A02186
R:Lopez de Castro, J.A.; Bragado, R.; Strong, D.M.; Strominger, J.L.
Biochemistry 22, 3961-3969, 1983
A:Title: Primary structure of papain-solubilized human histocompatibility antigen HLA
A:Reference number: A02186; MUID:84000412
A:Accession: A02186
A:Molecule type: protein
A:Residues: 1-270 <LOP>
C:Comment: Comparison of a number of class I allelic sequences (-B40 vs. -B7 and -A2
y restricted to the alpha-1 and alpha-2 domains. The most conspicuous clustering of v
ossible alloantigenic determinants of these antigens.
C:Genetics:
A:Gene: GDB:HLA-B
A:Cross-references: GDB:120048; OMIM:142830
A:Map position: 6p21.3-6p21.3
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology
C:Keywords: duplication; glycoprotein; heterodimer; membrane protein; transplantation
F:1-90/Domain: alpha-1 <EX1>
F:91-181/Domain: alpha-2 <EX2>
F:195-260/Domain: immunoglobulin homology <IMM>
F:86/Binding site: carbohydrate (Asn) (covalent) #status experimental
F:101-163,202-258/Disulfide bonds: #status predicted

```

```

Query Match 50.0%; Score 51; DB 1; Length 270;
Best Local Similarity 100.0%; Pred. No. 1.8; Indels 0; Gaps 0;
Matches 10; Conservative 0; Mismatches 0;

```

```

Qy 11 RESLNRLGY 20
|||||
Db 75 RESLNRLGY 84

```

```

RESULT 7
I68774
MHC HLA-B39 chain - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change 23-Jul-1999
C:Accession: I68774
R:Mueller, C.A.; Engler-Blum, G.; Gekeler, V.; Steiert, I.; Weiss, E.; Schmidt, H.
Immunogenetics 30, 200-207, 1989
A:Title: Genetic and serological heterogeneity of the supertypic HLA-B locus specific
A:Reference number: I54463; MUID:89379286
A:Accession: I68774
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-274 <RES>
A:Cross-references: GB:M29865; NID:g187676; PIDN:AAA36223.1; PID:g187677
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

```

```

Query Match 50.0%; Score 51; DB 2; Length 274;
Best Local Similarity 100.0%; Pred. No. 1.8;

```

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 RESLNRLGY 20

Db 75 RESLNRLGY 84

RESULT 8

S24439

Class I histocompatibility antigen HLA-B-3901 - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 22-Nov-1993 #sequence\_revision 01-Nov-1996 #text\_change 07-Nov-1997

C:Accession: S24439

R:Watkins, D.I.; McAdam, S.N.; Liu, X.; Strang, C.R.; Milford, E.L.; Levine, C.G.; Garbe

Nature 357, 329-333, 1992

A:Title: New recombinant HLA-B alleles in a tribe of South American Amerindians indicate

A:Reference number: S24027; MUID:92269956

A:Accession: S24439

A:Molecule type: mRNA

A:Residues: 1-274 <WAT>

C:Genetics:

A:Gene: HLA-B-3901

C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

C:Keywords: glycoprotein

F:1-274/Product: class I histocompatibility antigen HLA-B-3901 #status predicted <MAT>

F:1-90/Domain: alpha-1 #status predicted <EX1>

F:91-182/Domain: alpha-2 #status predicted <EX2>

F:196-261/Domain: immunoglobulin homology <IMX>

F:86/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:101-184,203-259/Disulfide bonds: #status predicted

Query Match

Best Local Similarity 50.0%; Score 51; DB 2; Length 274;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 RESLNRLGY 20

Db 75 RESLNRLGY 84

RESULT 9

I68701

cell surface antigen - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 07-Jun-1996 #sequence\_revision 07-Jun-1996 #text\_change 23-Jul-1999

C:Accession: I68701

R:Arnot, D.; Lillie, J.W.; Auffray, C.; Kappes, D.; Strominger, J.L.

Immunogenetics 20, 237-252, 1984

A:Title: Inter-locus and intra-allelic polymorphisms of HLA class I antigen gene mRNA.

A:Reference number: I54412; MUID:84287690

A:Accession: I68701

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-300 <RES>

A:Cross-references: GB:M27540; NID:g187733; PIDN:AAA59638.1; PID:g386890

C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

C:Keywords: surface antigen

Query Match

Best Local Similarity 50.0%; Score 51; DB 2; Length 300;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 RESLNRLGY 20

Db 37 RESLNRLGY 46

RESULT 10

I36956

MHC class II chain - chimpanzee (fragment)

C:Species: Pan troglodytes (chimpanzee)

C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 23-Jul-1999

C:Accession: I36956

R:Parham, P.; Lawlor, D.A.; Lomen, C.E.; Ennis, P.D.

J. Immunol. 142, 3937-3950, 1989

A:Title: Diversity and diversification of HLA-A,B,C alleles.

A:Reference number: I36956; MUID:892335215

A:Accession: I36956

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-308 <RES>

A:Cross-references: GB:M24044; NID:g176812; PIDN:AAA53423.1; PID:g176813

C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match

Best Local Similarity 50.0%; Score 51; DB 2; Length 308;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 RESLNRLGY 20

Db 45 RESLNRLGY 54

RESULT 11

I68747

MHC class I lymphocyte antigen - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 07-Jun-1996 #sequence\_revision 07-Jun-1996 #text\_change 23-Jul-1999

C:Accession: I68747

R:Pohl, H.; Kuon, W.; Tabaczewski, P.; Doerner, C.; Weiss, E.H.

Immunogenetics 29, 297-307, 1989

A:Title: Allelic variation in HLA-B and HLA-C sequences and the evolution of the HLA-

A:Reference number: I54457; MUID:89233295

A:Accession: I68747

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-350 <RES>

A:Cross-references: GB:M28204; NID:g576472; PIDN:AAA53257.1; PID:g576473

C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match

Best Local Similarity 50.0%; Score 51; DB 2; Length 350;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 RESLNRLGY 20

Db 87 RESLNRLGY 96

RESULT 12

I54308

MHC HLA B71 - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 23-Jul-1999

C:Accession: I54308

R:Rodriguez, S.G.; Johnson, A.H.; Hurley, C.K.

Hum. Immunol. 37, 192-194, 1993

A:Title: Molecular characterization of HLA-B71 from an African American individual.

A:Reference number: I54308; MUID:94064392

A:Accession: I54308

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-350 <RES>

A:Cross-references: GB:L07950; NID:g307236; PIDN:AAA59683.1; PID:g307237

C:Genetics:

A:Gene: GDB:HLA-B

A:Cross-references: GDB:I20048; OMIM:142830

A:Map position: 6p21.3-6p21.3

C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match

Best Local Similarity 50.0%; Score 51; DB 2; Length 350;

Best Local Similarity 100.0%; Pred. No. 2.3; DB 2: Length 354;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RESLRNLGY 20  
| | | | | | | | | |  
DB 99 RESLRNLGY 108

RESULT 13  
S24436  
class I histocompatibility antigen HLA-B-3504 precursor - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 22-Nov-1993 #sequence\_revision 18-Jul-1997 #text\_change 07-Nov-1997  
C:Accession: S24436  
R:Watkins, D.I.; McAdam, S.N.; Liu, X.; Strang, C.R.; Milford, E.L.; Levine, C.G.; Garber  
Nature 357, 329-333, 1992  
A:Title: New recombinant HLA-B alleles in a tribe of South American Amerindians indicate  
A:Reference number: S24027; MUID:92269956  
A:Accession: S24436  
A:Molecule type: mRNA  
A:Residues: 1-354 <WAT>  
C:Genetics:  
A:Gene: HLA-B-3504  
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology  
C:Keywords: glycoprotein; transmembrane protein  
F:1-17/Domain: signal sequence #status predicted <SIG>  
F:17-354/Product: class I histocompatibility antigen HLA-B-3504 #status predicted <WAT>  
F:17-299/Domain: extracellular #status predicted <EXT>  
F:107-198/Domain: alpha-2 #status predicted <EX2>  
F:212-277/Domain: immunoglobulin homology <IMM>  
F:300-323/Domain: transmembrane #status predicted <TM>  
F:324-354/Domain: intracellular #status predicted <INT>  
F:102/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:117-180,219-275/Disulfide bonds: #status predicted

Query Match 50.0%; Score 51; DB 2: Length 354;  
Best Local Similarity 100.0%; Pred. No. 2.4;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RESLRNLGY 20  
| | | | | | | | | |  
DB 91 RESLRNLGY 100

RESULT 14  
S24437  
class I histocompatibility antigen HLA-B-4802 precursor - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 22-Nov-1993 #sequence\_revision 18-Jul-1997 #text\_change 07-Nov-1997  
C:Accession: S24437  
R:Watkins, D.I.; McAdam, S.N.; Liu, X.; Strang, C.R.; Milford, E.L.; Levine, C.G.; Garber  
Nature 357, 329-333, 1992  
A:Title: New recombinant HLA-B alleles in a tribe of South American Amerindians indicate  
A:Reference number: S24027; MUID:92269956  
A:Accession: S24437  
A:Molecule type: mRNA  
A:Residues: 1-354 <WAT>  
C:Genetics:  
A:Gene: HLA-B-4802  
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology  
C:Keywords: glycoprotein; transmembrane protein  
F:1-17/Domain: signal sequence #status predicted <SIG>  
F:17-354/Product: class I histocompatibility antigen HLA-B-4802 #status predicted <WAT>  
F:17-299/Domain: extracellular #status predicted <EXT>  
F:107-198/Domain: alpha-2 #status predicted <EX2>  
F:212-277/Domain: immunoglobulin homology <IMM>  
F:300-323/Domain: transmembrane #status predicted <TM>  
F:324-354/Domain: intracellular #status predicted <INT>  
F:102/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:117-180,219-275/Disulfide bonds: #status predicted

Query Match 50.0%; Score 51; DB 2: Length 354;  
Best Local Similarity 100.0%; Pred. No. 2.4;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RESLRNLGY 20  
| | | | | | | | | |  
DB 91 RESLRNLGY 100

RESULT 15  
S24438  
class I histocompatibility antigen HLA-B-4801 precursor - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 22-Nov-1993 #sequence\_revision 18-Jul-1997 #text\_change 07-Nov-1997  
C:Accession: S24438  
R:Watkins, D.I.; McAdam, S.N.; Liu, X.; Strang, C.R.; Milford, E.L.; Levine, C.G.; Ga  
Nature 357, 329-333, 1992  
A:Title: New recombinant HLA-B alleles in a tribe of South American Amerindians indic  
A:Reference number: S24027; MUID:92269956  
A:Accession: S24438  
A:Molecule type: mRNA  
A:Residues: 1-354 <WAT>  
C:Genetics:  
A:Gene: HLA-B-4801  
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology  
C:Keywords: glycoprotein; transmembrane protein  
F:1-17/Domain: signal sequence #status predicted <SIG>  
F:17-354/Product: class I histocompatibility antigen HLA-B-4801 #status predicted <WA  
F:17-299/Domain: extracellular #status predicted <EXT>  
F:107-198/Domain: alpha-2 #status predicted <EX2>  
F:212-277/Domain: immunoglobulin homology <IMM>  
F:300-323/Domain: transmembrane #status predicted <TM>  
F:324-354/Domain: intracellular #status predicted <INT>  
F:102/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:117-180,219-275/Disulfide bonds: #status predicted

Query Match 50.0%; Score 51; DB 2: Length 354;  
Best Local Similarity 100.0%; Pred. No. 2.4;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RESLRNLGY 20  
| | | | | | | | | |  
DB 91 RESLRNLGY 100

Search completed: February 7, 2000, 18:04:38  
Job time: 22204 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2000, 01:25:59 ; Search time 68.63 Seconds  
(without alignments)  
8.703 Million cell updates/sec

Title: US-08-653-294-27

Perfect score: 102

Sequence: 1 YGRNLSRSLRLRGY 20

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 82229 seqs, 29864866 residues

Total number of hits satisfying chosen parameters: 82229

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : SwissProt\_38.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	51	50.0	270	1	1B33_HUMAN
2	51	50.0	354	1	1B24_HUMAN
3	51	50.0	362	1	1A02_PANTR
4	51	50.0	362	1	1B02_HUMAN
5	51	50.0	362	1	1B02_PANTR
6	51	50.0	362	1	1B04_HUMAN
7	51	50.0	362	1	1B07_HUMAN
8	51	50.0	362	1	1B08_HUMAN
9	51	50.0	362	1	1B10_HUMAN
10	51	50.0	362	1	1B11_HUMAN
11	51	50.0	362	1	1B12_HUMAN
12	51	50.0	362	1	1B13_HUMAN
13	51	50.0	362	1	1B21_HUMAN
14	51	50.0	362	1	1B22_HUMAN
15	51	50.0	362	1	1B23_HUMAN
16	51	50.0	362	1	1B25_HUMAN
17	51	50.0	362	1	1B26_HUMAN
18	51	50.0	362	1	1B27_HUMAN
19	51	50.0	362	1	1B28_HUMAN
20	51	50.0	362	1	1B31_HUMAN
21	51	50.0	362	1	1B32_HUMAN
22	51	50.0	362	1	1B34_HUMAN
23	51	50.0	362	1	1B35_HUMAN
24	51	50.0	362	1	1B36_HUMAN
25	51	50.0	362	1	1B38_HUMAN
26	51	50.0	362	1	1B39_HUMAN
27	51	50.0	362	1	1B43_HUMAN
28	51	50.0	362	1	1B46_HUMAN
29	51	50.0	362	1	1B48_HUMAN
30	51	50.0	362	1	1B55_HUMAN
31	51	50.0	362	1	1B56_HUMAN
32	51	50.0	362	1	1B57_HUMAN
33	51	50.0	362	1	1B58_HUMAN
34	51	50.0	362	1	1B59_HUMAN

35 51 50.0 362 1 1B63\_HUMAN P30498 homo sapien  
36 47.5 46.6 605 1 GLCA\_SOYBN P33916 glycine max  
37 46 45.1 346 1 HRPJ\_PSESY Q05395 pseudomonas  
38 45.5 44.6 366 1 1C01\_HUMAN P30499 homo sapien  
39 44.5 43.6 639 1 GLCX\_SOYBN P11827 glycine max  
40 44 43.1 342 1 1CXX\_HUMAN P10321 homo sapien  
41 44 43.1 362 1 1B44\_HUMAN P30484 homo sapien  
42 44 43.1 366 1 1C01\_PANTR P30686 pan troglod  
43 44 43.1 366 1 1C02\_HUMAN P30500 homo sapien  
44 44 43.1 366 1 1C05\_HUMAN P04222 homo sapien  
45 44 43.1 366 1 1C06\_HUMAN P30503 homo sapien

## ALIGNMENTS

RESULT 1  
ID 1B33\_HUMAN STANDARD; PRT; 270 AA.  
AC P01890;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 01-APR-1993 (Rel. 25, Last annotation update)  
DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, BW-60(B-40) B\*4001 ALPHA CHAIN  
DE (FRAGMENT).  
GN HLA-B OR HLAB.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE.  
RX MEDLINE; 84000412.  
RA LOPEZ DE CASTRO J.A., BRAGADO R., STRONG D.M., STROMINGER J.L.;  
RT "Primary structure of papain-solubilized human histocompatibility  
RT antigen HLA-B\*40 (-B\*60). An outline of alloantigenic determinants."  
RL Biochemistry 22:3961-3969(1983).  
CC -1- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO  
CC THE IMMUNE SYSTEM.  
CC -1- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
CC MICROGLOBULIN).  
DR PIR; A02186; HLU40.  
DR HSP; P30460; IAGB.  
DR MIM; 142830; .  
DR PROSITE; PS00290; IG\_MHC; 1.  
DR PFAM; PF00047; Ig; 1.  
DR PFAM; PF00129; MHC\_I; 1.  
KW MHC I; Glycoprotein.  
FT DOMAIN 1 90 EXTRACELLULAR ALPHA-1.  
FT DOMAIN 91 181 EXTRACELLULAR ALPHA-2.  
FT DOMAIN 182 >270 EXTRACELLULAR ALPHA-3.  
FT CARBOHYD 86 86  
FT DISULFID 101 163  
FT DISULFID 202 258  
FT NON\_TER 270 270  
SQ SEQUENCE 270 AA; BFE44EFF CRC32;

Query Match 50.08; Score 51; DB 1; Length 270;  
Best Local Similarity 100.0%; Pred. No. 0.59; Indels 0; Gaps 0;  
Matches 10; Conservative 0; Mismatches 0;  
Qy 11 RESLRNLRGY 20  
Db 75 RESLRNLRGY 84

RESULT 2  
ID 1B24\_HUMAN STANDARD; PRT; 354 AA.  
AC P30470;  
DT 01-APR-1993 (Rel. 25, Created)  
DT 01-APR-1993 (Rel. 25, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-35 B\*3504 ALPHA CHAIN  
DE PRECURSOR (FRAGMENT).

GN HLA-B OR HLAB.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

OC Eutheria; Primates; Catarrhini; Homidae; Homo.

RN [1]

RP SEQUENCE FROM N.A.

RA MEDLINE; 92269556.

RA WATKINS D.I., MCADAM S.N., LIU X., STANG C.R., MILFORD E.L.,

RA LEVINE C.G., GARBER T.L., DOGAN A.L., LORD C.I., GHIM S.H.,

RA TROUP G.M., HUGHES A.L., LETVIN N.L.;

RT "New recombinant HLA-B alleles in a tribe of South American

RT Amerindians indicate rapid evolution of MHC class I loci.;

RL Nature 357:329-333(1992).

CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO

CC THE IMMUNE SYSTEM.

CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-

CC MICROGLOBULIN).

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL; M86403; ; NOT\_ANNOTATED\_CDS.

CC HSSP; P30685; 1AIN.

CC MIM; 142830; ;

CC PROSITE; PS00290; IG\_MHC; 1.

CC PFAM; PF00047; ig; 1.

CC PFAM; PF00129; MHC\_I; 1.

CC MHC I; Transmembrane; Glycoprotein; Signal.

CC NON\_TER 1 1

CC SIGNAL <1 16

CC CHAIN 17 354

CC HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,

CC B-35 B\*3504 ALPHA CHAIN.

CC EXTRACELLULAR ALPHA-1.

CC EXTRACELLULAR ALPHA-2.

CC EXTRACELLULAR ALPHA-3.

CC CONNECTING PEPTIDE.

CC CYTOPLASMIC TAIL.

CC BY SIMILARITY.

CC BY SIMILARITY.

CC BY SIMILARITY.

CC CARBOHYD 102 102

CC DISULFID 117 180

CC SEQUENCE 354 AA; 39617 MW; 6564795A CRC32;

CC -----

Query Match 50.0%; Score 51; DB 1; Length 354;

Best Local Similarity 100.0%; Pred. No. 0.8;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RESLRNLCY 20

Db 91 RESLRNLCY 100

|||||

RESULT 3

1A02\_PANTR STANDARD; PRT; 362 AA.

AC P16210;

DT 01-APR-1990 (Rel. 14, Created)

DT 01-APR-1990 (Rel. 14, Last sequence update)

DT 01-APR-1993 (Rel. 25, Last annotation update)

DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, A-5 ALPHA CHAIN PRECURSOR.

OS Pan troglodytes (Chimpanzee).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

OC Eutheria; Primates; Catarrhini; Homidae; Pan.

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE; 90201944.  
RA LAWOR D.A., WARREN E., WARD F.E., PARHAM P.;  
RT "Comparison of class I MHC alleles in humans and apes.";  
RL Immunol. Rev. 113:147-185(1990).  
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO  
CC THE IMMUNE SYSTEM.  
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
CC MICROGLOBULIN).

CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL; M30679; AAA87971.1; .

CC HSSP; P30685; 1AIN.

CC PROSITE; PS00290; IG\_MHC; 1.

CC PFAM; PF00047; ig; 1.

CC PFAM; PF00129; MHC\_I; 1.

CC MHC I; Transmembrane; Glycoprotein; Signal.

CC SIGNAL 1 24

CC CHAIN 25 362

CC HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,

CC A-5 ALPHA CHAIN.

CC EXTRACELLULAR ALPHA-1.

CC EXTRACELLULAR ALPHA-2.

CC EXTRACELLULAR ALPHA-3.

CC CONNECTING PEPTIDE.

CC CYTOPLASMIC TAIL.

CC BY SIMILARITY.

CC BY SIMILARITY.

CC BY SIMILARITY.

CC CARBOHYD 110 110

CC SEQUENCE 362 AA; 40487 MW; 97E6CE8A CRC32;

CC -----

Query Match 50.0%; Score 51; DB 1; Length 362;

Best Local Similarity 100.0%; Pred. No. 0.8;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RESLRNLCY 20

Db 99 RESLRNLCY 108

|||||

RESULT 4

1B02\_HUMAN STANDARD; PRT; 362 AA.

AC P01889;

DT 21-JUL-1986 (Rel. 01, Created)

DT 01-FEB-1991 (Rel. 17, Last sequence update)

DT 01-OCT-1996 (Rel. 34, Last annotation update)

DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-7 B\*0702 ALPHA CHAIN

DE PRECURSOR (B7.2).

GN HLA-B OR HLAB.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

OC Eutheria; Primates; Catarrhini; Homidae; Homo.

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE; 90207291.

RA ENNIS P.D., ZEMMOUR J., SALTER R.D., PARHAM P.;

RT "Rapid cloning of HLA-A,B cDNA by using the polymerase chain

RT reaction: frequency and nature of errors produced in amplification.";

RL Proc. Natl. Acad. Sci. U.S.A. 87:2833-2837(1990).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE; 90315860.

RA PARHAM P., BENJAMIN R.J., CHEN B.P., CLAYBERGER C., ENNIS P.D.,

RA KRENSKY A.M., LAWOR D.A., LITTMAN D.R., NORRMENT A.M., ORR H.T.,

RA SALTER R.D., ZEMMOUR J.;



"Diversity of class I HLA molecules: functional and evolutionary interactions with T cells.";  
 RL Cold Spring Harb. Symp. Quant. Biol. 54:529-543(1989).  
 (3)  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE: 85287366.  
 RX SOOD A.K., PAN J., BIRO P.A., PEREIRA D., SRIVASTAVA R., REDDY V.B.,  
 RA DUCEMAN B.W., WEISSMAN S.M.;  
 RT "Structure and polymorphism of class I MHC antigen mRNA.";  
 RL Immunogenetics 22:101-121(1985).  
 (4)  
 RN SEQUENCE FROM N.A.  
 RP ELLEXSON M.E., ZHANG L., HILDEBRAND W.H.;  
 RA Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.  
 (5)  
 RN SEQUENCE OF 25-295.  
 RX MEDLINE: 80088278.  
 RA ORR H.T., LOPEZ DE CASTRO J.A., LANCET D., STROMINGER J.L.;  
 RT "Complete amino acid sequence of a papain-solubilized human  
 RT histocompatibility antigen, HLA-B\*7.2. Sequence determination and  
 search for homologies.";  
 RL Biochemistry 18:5711-5720(1979).  
 CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO  
 CC THE IMMUNE SYSTEM.  
 CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
 CC MICROGLOBULIN).  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M32317; AAA36230.1; -;  
 DR EMBL: M16102; AAA59622.1; ALT\_SEQ.  
 DR EMBL: U29057; AAA91229.1; -;  
 DR PIR: A02185; HLHUB7.  
 DR PIR: B35997; B35997.  
 DR HSSP: P30460; IAGB.  
 DR MIM: 142830; -;  
 DR PROSITE: PS00290; IG\_MHC; 1.  
 DR PFAM: PF00047; ig; 1.  
 DR PFAM: PF00129; MHC I; 1.  
 DR MHC I; Transmembrane; Glycoprotein; Signal.  
 KW SIGNAL 1 24  
 FT CHAIN 25 362 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,  
 FT B-7 B\*0702 ALPHA CHAIN.  
 FT DOMAIN 25 114  
 FT DOMAIN 115 206 EXTRACELLULAR ALPHA-1.  
 FT DOMAIN 207 298 EXTRACELLULAR ALPHA-2.  
 FT DOMAIN 299 309 EXTRACELLULAR ALPHA-3.  
 FT DOMAIN 310 333 CONNECTING PEPTIDE.  
 FT TRANSMEM 334 362 CYTOPLASMIC TAIL.  
 FT CARBOHYD 110 110  
 FT DISULFID 125 188  
 FT DISULFID 227 283  
 FT CONFLICT 15 18 AALA -> GPW (IN REF. 3).  
 FT CONFLICT 266 266 Q -> E (IN REF. 5).  
 FT CONFLICT 268 268 W -> S (IN REF. 3).  
 FT CONFLICT 297 297 R -> G (IN REF. 3).  
 FT CONFLICT 314 315 GL -> RP (IN REF. 3).  
 FT CONFLICT 362 AA; 40460 MW; 87B2ED84 CRC32;  
 SQ SEQUENCE 362 AA; 40460 MW; 87B2ED84 CRC32;  
 Query Match 50.0%; Score 51; DB 1; Length 362;  
 Best Local Similarity 100.0%; Pred. No. 0.82;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 11 RESRLNLRGY 20  
 DB 99 RESRLNLRGY 108  
 RESULT 6  
 ID 1B04\_HUMAN STANDARD; PRT; 362 AA.  
 AC P30460;

RESULT 5  
 ID 1B02\_PANTR STANDARD; PRT; 362 AA.  
 AC P13751;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 01-APR-1993 (Rel. 25, Last annotation update)  
 DE CHLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-2 ALPHA CHAIN PRECURSOR.  
 OS Pan troglodytes (Chimpanzee).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Pan.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 89030641.  
 RA MAYER W.E., JONKER M., KLEIN D., IVANYI P., VAN SEVENTER G.,  
 RA KLEIN J.;  
 RT "Nucleotide sequences of chimpanzee MHC class I alleles: evidence for  
 RT trans-species mode of evolution.";  
 RL EMBO J. 7:2765-2774(1988).  
 RN [2]  
 RP REVISIONS.  
 RA MAYER W.;  
 RL Submitted (FEB-1989) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO  
 CC THE IMMUNE SYSTEM.  
 CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
 CC MICROGLOBULIN).  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X13116; CAA31508.1; -;  
 DR PIR: S03538; S03538.  
 DR HSSP: P30685; IAIN.  
 DR PROSITE: PS00290; IG\_MHC; 1.  
 DR PFAM: PF00047; ig; 1.  
 DR PFAM: PF00129; MHC I; 1.  
 DR MHC I; Transmembrane; Glycoprotein; Signal.  
 KW SIGNAL 1 24  
 FT CHAIN 25 362 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,  
 FT B-2 ALPHA CHAIN.  
 FT DOMAIN 25 114  
 FT DOMAIN 115 206 EXTRACELLULAR ALPHA-1.  
 FT DOMAIN 207 298 EXTRACELLULAR ALPHA-2.  
 FT DOMAIN 299 308 EXTRACELLULAR ALPHA-3.  
 FT TRANSMEM 309 332 CONNECTING PEPTIDE.  
 FT DOMAIN 333 362 CYTOPLASMIC TAIL.  
 FT DISULFID 125 188 BY SIMILARITY.  
 FT DISULFID 227 283 BY SIMILARITY.  
 FT CARBOHYD 110 110 BY SIMILARITY.  
 FT SEQUENCE 362 AA; 40488 MW; 4BF65A6C CRC32;  
 Query Match 50.0%; Score 51; DB 1; Length 362;  
 Best Local Similarity 100.0%; Pred. No. 0.82;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 11 RESRLNLRGY 20  
 DB 99 RESRLNLRGY 108  
 RESULT 6  
 ID 1B04\_HUMAN STANDARD; PRT; 362 AA.  
 AC P30460;

```

DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-8 B*0801 ALPHA CHAIN
DE PRECURSOR.
GN HLA-B OR HLAB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 89235215.
RA PARHAM P., LAWLER D.A., LOMEN C.E., ENNIS P.D.;
RT "Diversity and diversification of HLA-A,B,C alleles.";
RL J. Immunol. 142:3937-3950(1989).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 25-300.
RX MEDLINE; 97130420.
RA REID S.W., MCADAM S., SMITH K.J., KLENERMAN P., O'CALLAGHAN C.A.,
RA HARLOS K., JAKOBSEN B.K., MCMICHAEL A.J., BELL J.I., STUART D.I.,
RA JONES E.Y.;
RT "Antagonist HIV-1 Gag peptides induce structural changes in HLA B8.";
RL J. Exp. Med. 184:2279-2286(1996).
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC THE IMMUNE SYSTEM.
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC MICROGLOBULIN).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; M24036; AAA52662.1; -.
DR PDB; 1AGB; 16-JUN-97.
DR PDB; 1AGC; 16-JUN-97.
DR PDB; 1AGD; 16-JUN-97.
DR PDB; 1AGE; 16-JUN-97.
DR PDB; 1AGF; 16-JUN-97.
DR MIM; 142830; -.
DR PROSITE; PS00290; IG_MHC; 1.
DR PFAM; PF00047; ig; 1.
DR PFAM; PF00129; MHC_I; 1.
KW MHC I; Transmembrane; Glycoprotein; Signal; 3D-structure.
FT SIGNAL 1 24
FT CHAIN 25 362
FT DOMAIN 25 114
FT DOMAIN 115 206
FT DOMAIN 207 298
FT DOMAIN 299 309
FT TRANSMEM 310 333
FT DOMAIN 334 362
FT CARBOHYD 110 110
FT DISULFID 125 188
FT DISULFID 227 283
SQ SEQUENCE 362 AA; 40331 MW; 1467B9EB CRC32;

Query Match 50.0%; Score 51; DB 1; Length 362;
Best Local Similarity 100.0%; Pred. No. 0.82;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RESLNRLRGY 20
DB 99 RESLNRLRGY 108

RESULT 7
1B07_HUMAN

```

```

ID 1B07_HUMAN STANDARD; PRT; 362 AA.
AC P30462;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-APR-1993 (Rel. 25, Last annotation update)
DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-14 B*1401 ALPHA CHAIN
DE PRECURSOR.
GN HLA-B OR HLAB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 89235215.
RA PARHAM P., LAWLER D.A., LOMEN C.E., ENNIS P.D.;
RT "Diversity and diversification of HLA-A,B,C alleles.";
RL J. Immunol. 142:3937-3950(1989).
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC THE IMMUNE SYSTEM.
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC MICROGLOBULIN).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; M24040; AAA59661.1; -.
DR HSSP; P30460; 1AGB.
DR MIM; 142830; -.
DR PROSITE; PS00290; IG_MHC; 1.
DR PFAM; PF00047; ig; 1.
DR PFAM; PF00129; MHC_I; 1.
KW MHC I; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 24
FT CHAIN 25 362
FT DOMAIN 25 114
FT DOMAIN 115 206
FT DOMAIN 207 298
FT DOMAIN 299 309
FT TRANSMEM 310 333
FT DOMAIN 334 362
FT CARBOHYD 110 110
FT DISULFID 125 188
FT DISULFID 227 283
SQ SEQUENCE 362 AA; 40358 MW; 9BED8199 CRC32;

Query Match 50.0%; Score 51; DB 1; Length 362;
Best Local Similarity 100.0%; Pred. No. 0.82;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RESLNRLRGY 20
DB 99 RESLNRLRGY 108

RESULT 8
1B08_HUMAN
ID 1B08_HUMAN STANDARD; PRT; 362 AA.
AC P30463;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-APR-1993 (Rel. 25, Last annotation update)
DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, BW-65(B-14) B*1402 ALPHA CHAIN
DE PRECURSOR.
GN HLA-B OR HLAB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

```

[illegible]

RESULT	10
1B1L HUMAN	
ID ID	HUMAN
AC AC	STANDARD;
P30465;	PRT; 362 AA.
01-APR-1993	(Rel. 25, Created)
01-APR-1993	(Rel. 25, Last sequence update)
15-JUL-1998	(Rel. 38, Last annotation update)
HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,	BW*72(BW-70) B*1503 ALPHA
CHAIN PRECURSOR.	
HLA-B OR HLAB.	
GN GN	
OS Homo sapiens (Human).	
OC Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;	
Eutheria; Primates; Catarrhini; Hominoidea; Homo.	
[1]	
RN SEQUENCE FROM N.A.	
RP RP	MEDLINE: 93056508.
RX RX	
RA MADRIGNAL J.A., BELICH M.P., HILDEBRAND W.H., BENJAMIN R.J.,	
LITTLE A.M., ZEMMOUR J., ENNIS P.D., WARD F.E., PETZL-EKLER M.L.,	
MARTELL R.W., DU TOIT E.D., PARHAM P.;	
"Distinctive HLA-A,B antigens of black populations formed by	
interallelic conversion.";	
RT RT	

```

RL J. Immunol. 149:3411-3415(1992).
CC -1- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC THE IMMUNE SYSTEM.
CC -1- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC MICROGLOBULIN).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X61709; CAA43878.1; -
DR PIR; S16789; S16789.
DR HSP; P30685; 1A1N.
DR MIM; 142830; -
DR PROSITE; PS00290; IG_MHC; 1.
DR PFAM; PF00047; Ig; 1.
DR PFAM; PF00129; MHC_I; 1.
KW MHC I; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 24
FT CHAIN 25 362 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
FT BW-72(BW-70) B*1503 ALPHA CHAIN.
FT DOMAIN 25 114 EXTRACELLULAR ALPHA-1.
FT DOMAIN 115 206 EXTRACELLULAR ALPHA-2.
FT DOMAIN 207 298 EXTRACELLULAR ALPHA-3.
FT DOMAIN 299 309 CONNECTING PEPTIDE.
FT TRANSMEM 310 333
FT DOMAIN 334 362 CYTOPLASMIC TAIL.
FT CARBOHYD 110 110 BY SIMILARITY.
FT DISULFID 125 188 BY SIMILARITY.
FT DISULFID 227 283 BY SIMILARITY.
FT SEQUENCE 362 AA; 40387 MW; 99D70546 CRC32;
SQ

```

```

Query Match 50.0%; Score 51; DB 1; Length 362;
Best Local Similarity 100.0%; Pred. No. 0.82;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 11 RESLRNLRGY 20
| | | | | | | |
DB 99 RESLRNLRGY 108

```

```

RESULT 11
1B12_HUMAN
ID 1B12_HUMAN STANDARD; PRT; 362 AA.
AC P30513;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-62 B*1504 ALPHA CHAIN
DE PRECURSOR.
DE HLA-B OR HLAB.
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominoidea; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 92269956.
RA WATKINS D.I., MCADAM S.N., LIU X., STANG C.R., MILFORD E.L.,
RA LEVINE C.G., GABER T.L., DOGON A.L., LORD C.I., GHIM S.H.,
RA TROUP G.M., HUGHES A.L., LETVIN N.L.;
RT "New recombinant HLA-B alleles in a tribe of South American
RT Amerindians indicate rapid evolution of MHC class I loci."
RL Nature 357:329-333(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-BLOOD;
RA RAWOS M., BARBER D.F., LAYRISSE Z., DE CASTRO J.A.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.

```

```

CC -1- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC THE IMMUNE SYSTEM.
CC -1- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC MICROGLOBULIN).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M84382; AAA59632.1; -
DR EMBL; U70528; AAB16918.1; -
DR PIR; S24433; S24433.
DR HSP; P30685; 1A1N.
DR MIM; 142830; -
DR PROSITE; PS00290; IG_MHC; 1.
DR PFAM; PF00047; Ig; 1.
DR PFAM; PF00129; MHC_I; 1.
KW MHC I; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 24
FT CHAIN 25 362 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
FT B-62 B*1504 ALPHA CHAIN.
FT DOMAIN 25 114 EXTRACELLULAR ALPHA-1.
FT DOMAIN 115 206 EXTRACELLULAR ALPHA-2.
FT DOMAIN 207 298 EXTRACELLULAR ALPHA-3.
FT DOMAIN 299 309 CONNECTING PEPTIDE.
FT TRANSMEM 310 333
FT DOMAIN 334 362 CYTOPLASMIC TAIL.
FT CARBOHYD 110 110 BY SIMILARITY.
FT DISULFID 125 188 BY SIMILARITY.
FT DISULFID 227 283 BY SIMILARITY.
FT SEQUENCE 362 AA; 40406 MW; E80FC24C CRC32;
SQ

```

```

Query Match 50.0%; Score 51; DB 1; Length 362;
Best Local Similarity 100.0%; Pred. No. 0.82;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 11 RESLRNLRGY 20
| | | | | | | |
DB 99 RESLRNLRGY 108

```

```

RESULT 12
1B13_HUMAN
ID 1B13_HUMAN STANDARD; PRT; 362 AA.
AC P30466;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-18 B*1801 ALPHA CHAIN
DE PRECURSOR.
DE HLA-B OR HLAB.
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominoidea; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 89235215.
RA PARHAM P., LAWLOR D.A., LOMEN C.E., ENNIS P.D.;
RT "Diversity and diversification of HLA-A,B,C alleles."
RL J. Immunol. 142:3937-3950(1989).
CC -1- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC THE IMMUNE SYSTEM.
CC -1- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC MICROGLOBULIN).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

3 (unavailable)



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2000, 19:16:13 ; Search time 176.54 Seconds  
(without alignments)  
7.855 Million cell updates/sec

Title: US-08-653-294-27  
Perfect score: 102  
Sequence: 1 YGRLNLSERRESLNLRG 20

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 225878 seqs, 69334122 residues

Total number of hits satisfying chosen parameters: 225878

Minimum DB seq length: 0  
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database : SPTRMBL12.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phase.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	51	50.0	81	7 O19523	O19523 homo sapien
2	51	50.0	81	7 O19525	O19525 homo sapien
3	51	50.0	81	7 O19527	O19527 homo sapien
4	51	50.0	81	7 O19529	O19529 homo sapien
5	51	50.0	81	7 O19531	O19531 homo sapien
6	51	50.0	81	7 O19533	O19533 homo sapien
7	51	50.0	81	7 O19535	O19535 homo sapien
8	51	50.0	81	7 O19537	O19537 homo sapien
9	51	50.0	83	7 O19539	O19539 homo sapien
10	51	50.0	83	7 O19541	O19541 homo sapien
11	51	50.0	83	7 O19543	O19543 homo sapien
12	51	50.0	89	4 Q3F452	Q3F452 homo sapien
13	51	50.0	89	7 P79615	P79615 homo sapien
14	51	50.0	89	7 P79620	P79620 homo sapien
15	51	50.0	89	7 O19548	O19548 homo sapien
16	51	50.0	89	7 O19549	O19549 homo sapien
17	51	50.0	89	7 O19550	O19550 homo sapien
18	51	50.0	89	7 O19551	O19551 homo sapien
19	51	50.0	89	7 O19640	O19640 homo sapien
20	51	50.0	89	7 O77967	O77967 homo sapien

21	51	50.0	89	7 O78174	O78174 homo sapien
22	51	50.0	89	7 O19638	O19638 homo sapien
23	51	50.0	89	7 P79487	P79487 homo sapien
24	51	50.0	89	7 O19587	O19587 homo sapien
25	51	50.0	89	7 O19576	O19576 homo sapien
26	51	50.0	89	7 O19577	O19577 homo sapien
27	51	50.0	89	7 O19581	O19581 homo sapien
28	51	50.0	89	7 O19583	O19583 homo sapien
29	51	50.0	112	7 O92671	O92671 homo sapien
30	51	50.0	134	7 Q29677	Q29677 homo sapien
31	51	50.0	137	7 Q29660	Q29660 homo sapien
32	51	50.0	137	7 Q29663	Q29663 homo sapien
33	51	50.0	137	7 Q29665	Q29665 homo sapien
34	51	50.0	137	7 Q29664	Q29664 homo sapien
35	51	50.0	147	7 Q29707	Q29707 homo sapien
36	51	50.0	165	7 P79665	P79665 homo sapien
37	51	50.0	165	7 P79666	P79666 homo sapien
38	51	50.0	166	7 O19777	O19777 homo sapien
39	51	50.0	166	7 O19778	O19778 homo sapien
40	51	50.0	172	7 O19776	O19776 homo sapien
41	51	50.0	176	7 Q29662	Q29662 homo sapien
42	51	50.0	179	7 Q29676	Q29676 homo sapien
43	51	50.0	180	7 Q29678	Q29678 homo sapien
44	51	50.0	180	7 O77960	O77960 homo sapien
45	51	50.0	181	7 Q29716	Q29716 homo sapien

ALIGNMENTS

RESULT 1  
O19523 PRELIMINARY; PRT; 81 AA.  
AC O19523;  
DT 01-JAN-1998 (TRENBLrel. 05, Created)  
DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)  
DT 01-NOV-1998 (TRENBLrel. 08, Last annotation update)  
DE MHC CLASS I ANTIGEN HLA-B (FRAGMENT).  
GN HLA-B.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA CHANDANAYINGTONG D., SIRIKONG M., LUANGTRAKOOL K., SRINAK D.,  
RA RUNGRONG E., BECHANDRA S.;  
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF014769; AAB67807.1;  
DR PFAM; PF00129; MHC\_I; 1.  
KW MHC.  
FT NON\_TER 1 1  
FT NON\_TER 81 81  
SQ SEQUENCE 81 AA; 9405 MW; 073087CE CRC32;  
  
Query Match 50.0%; Score 51; DB 7; Length 81;  
Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 11 RESLNLRGY 20  
Db 66 RESLNLRGY 75  
  
RESULT 2  
O19525 PRELIMINARY; PRT; 81 AA.  
ID O19525;  
AC O19525;  
DT 01-JAN-1998 (TRENBLrel. 05, Created)  
DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)  
DT 01-NOV-1998 (TRENBLrel. 08, Last annotation update)  
DE MHC CLASS I ANTIGEN HLA-B (FRAGMENT).  
GN HLA-B.

OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 [1]  
 RN SEQUENCE FROM N.A.  
 RA CHANDANAYINGYONG D., SIRIKONG M., LUANGTRAKOOL K., SRINAK D.,  
 RA RUNGRONG E., BEJCHANDRA S.;  
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF014771; AAB67809.1; -  
 DR PFAM: PF00129; MHC\_I; 1.  
 KW MHC.  
 FT NON\_TER 1  
 FT NON\_TER 81  
 SQ SEQUENCE 81 AA; 9405 MW; 073087CE CRC32;

Query Match 50.0%; Score 51; DB 7; Length 81;  
 Best Local Similarity 100.0%; Pred. No. 1.2;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RESLRNLRGY 20  
 DB 66 RESLRNLRGY 75

RESULT 3  
 ID O19527 PRELIMINARY; PRT; 81 AA.  
 AC O19527;  
 DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
 DE MHC CLASS I ANTIGEN HLA-B (FRAGMENT).  
 GN HLA-B.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA CHANDANAYINGYONG D., SIRIKONG M., LUANGTRAKOOL K., SRINAK D.,  
 RA RUNGRONG E., BEJCHANDRA S.;  
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF014773; AAB67811.1; -  
 DR PFAM: PF00129; MHC\_I; 1.  
 KW MHC.  
 FT NON\_TER 1  
 FT NON\_TER 81  
 SQ SEQUENCE 81 AA; 9405 MW; 073087CE CRC32;

Query Match 50.0%; Score 51; DB 7; Length 81;  
 Best Local Similarity 100.0%; Pred. No. 1.2;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RESLRNLRGY 20  
 DB 66 RESLRNLRGY 75

RESULT 4  
 ID O19529 PRELIMINARY; PRT; 81 AA.  
 AC O19529;  
 DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
 DE MHC CLASS I ANTIGEN HLA-B (FRAGMENT).  
 GN HLA-B.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.

RA CHANDANAYINGYONG D., SIRIKONG M., LUANGTRAKOOL K., SRINAK D.,  
 RA RUNGRONG E., BEJCHANDRA S.;  
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF014775; AAB67813.1; -  
 DR PFAM: PF00129; MHC\_I; 1.  
 KW MHC.  
 FT NON\_TER 1  
 FT NON\_TER 81  
 SQ SEQUENCE 81 AA; 9405 MW; 073087CE CRC32;

Query Match 50.0%; Score 51; DB 7; Length 81;  
 Best Local Similarity 100.0%; Pred. No. 1.2;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RESLRNLRGY 20  
 DB 66 RESLRNLRGY 75

RESULT 5  
 ID O19531 PRELIMINARY; PRT; 81 AA.  
 AC O19531;  
 DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
 DE MHC CLASS I ANTIGEN HLA-B (FRAGMENT).  
 GN HLA-B.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA CHANDANAYINGYONG D., SIRIKONG M., LUANGTRAKOOL K., SRINAK D.,  
 RA RUNGRONG E., BEJCHANDRA S.;  
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF014777; AAB67815.1; -  
 DR PFAM: PF00129; MHC\_I; 1.  
 KW MHC.  
 FT NON\_TER 1  
 FT NON\_TER 81  
 SQ SEQUENCE 81 AA; 9405 MW; 073087CE CRC32;

Query Match 50.0%; Score 51; DB 7; Length 81;  
 Best Local Similarity 100.0%; Pred. No. 1.2;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RESLRNLRGY 20  
 DB 66 RESLRNLRGY 75

RESULT 6  
 ID O19533 PRELIMINARY; PRT; 81 AA.  
 AC O19533;  
 DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
 DE MHC CLASS I ANTIGEN HLA-B (FRAGMENT).  
 GN HLA-B.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA CHANDANAYINGYONG D., SIRIKONG M., LUANGTRAKOOL K., SRINAK D.,  
 RA RUNGRONG E., BEJCHANDRA S.;  
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF014779; AAB67817.1; -  
 DR PFAM: PF00129; MHC\_I; 1.



KW MHC.  
FT NON\_TER 1  
FT NON\_TER 81  
SQ SEQUENCE 81 AA; 9405 MW; 073087CE CRC32;

Query Match 50.0%; Score 51; DB 7; Length 81;  
Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 RESLNLRGY 20  
Db 66 RESLNLRGY 75

RESULT 7  
O19535 PRELIMINARY; PRT; 81 AA.  
AC O19535;  
DT 01-JAN-1998 (TREMBlrel. 05, Created)  
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)  
DE MHC CLASS I ANTIGEN HLA-B (FRAGMENT).  
GN HLA-B.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA CHANDANAYINGYONG D., SIRIKONG M., LUANGTRAKOOL K., SRINAK D.,  
RA RUNGROUNG E., BEJCHANDRA S.;  
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF014781; AAB67819.1; -  
DR PFAM; PF00129; MHC\_I; 1.  
KW MHC.  
FT NON\_TER 1  
FT NON\_TER 81  
SQ SEQUENCE 81 AA; 9405 MW; 073087CE CRC32;

Query Match 50.0%; Score 51; DB 7; Length 81;  
Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 RESLNLRGY 20  
Db 66 RESLNLRGY 75

RESULT 8  
O19537 PRELIMINARY; PRT; 81 AA.  
AC O19537;  
DT 01-JAN-1998 (TREMBlrel. 05, Created)  
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)  
DE MHC CLASS I ANTIGEN HLA-B (FRAGMENT).  
GN HLA-B.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA CHANDANAYINGYONG D., SIRIKONG M., LUANGTRAKOOL K., SRINAK D.,  
RA RUNGROUNG E., BEJCHANDRA S.;  
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF014783; AAB67821.1; -  
DR PFAM; PF00129; MHC\_I; 1.  
KW MHC.  
FT NON\_TER 1  
FT NON\_TER 81  
SQ SEQUENCE 81 AA; 9405 MW; 073087CE CRC32;

Query Match 50.0%; Score 51; DB 7; Length 81;  
Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 RESLNLRGY 20  
Db 66 RESLNLRGY 75

RESULT 9  
O19539 PRELIMINARY; PRT; 83 AA.  
AC O19539;  
DT 01-JAN-1998 (TREMBlrel. 05, Created)  
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)  
DE MHC CLASS I ANTIGEN HLA-B (FRAGMENT).  
GN HLA-B.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA CHANDANAYINGYONG D., SIRIKONG M., LUANGTRAKOOL K., SRINAK D.,  
RA RUNGROUNG E., BEJCHANDRA S.;  
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF014785; AAB67823.1; -  
DR PFAM; PF00129; MHC\_I; 1.  
KW MHC.  
FT NON\_TER 1  
FT NON\_TER 83  
SQ SEQUENCE 83 AA; 9731 MW; 24B8D666 CRC32;

Query Match 50.0%; Score 51; DB 7; Length 83;  
Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 RESLNLRGY 20  
Db 58 RESLNLRGY 77

RESULT 10  
O19541 PRELIMINARY; PRT; 83 AA.  
AC O19541;  
DT 01-JAN-1998 (TREMBlrel. 05, Created)  
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)  
DE MHC CLASS I ANTIGEN HLA-B (FRAGMENT).  
GN HLA-B.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA CHANDANAYINGYONG D., SIRIKONG M., LUANGTRAKOOL K., SRINAK D.,  
RA RUNGROUNG E., BEJCHANDRA S.;  
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF014787; AAB67825.1; -  
DR PFAM; PF00129; MHC\_I; 1.  
KW MHC.  
FT NON\_TER 1  
FT NON\_TER 83  
SQ SEQUENCE 83 AA; 9731 MW; 24B8D666 CRC32;

Query Match 50.0%; Score 51; DB 7; Length 83;  
Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 RESLRNLRGY 20  
 |||||  
 Db 68 RESLRNLRGY 77

## RESULT 11

O19543 PRELIMINARY; PRT; 83 AA.  
 AC O19543;  
 DT 01-JAN-1998 (TREMBlrel. 05, Created)  
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)  
 DE MHC CLASS I ANTIGEN HLA-B (FRAGMENT).  
 GN HLA-B.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA CHANDANAYONG D., SIRIKONG M., LUANGTRAKOOL K., SRINAK D.,  
 RA RUNGROUNG E., BEJCHANDRA S.,  
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF014789; AAB67827.1; -.  
 DR PFAM; PF00129; MHC\_I; 1.  
 KW MHC.  
 FT NON\_TER 1 1  
 FT NON\_TER 83 83  
 SQ SEQUENCE 83 AA; 9731 MW; 248D666 CRC32;

Query Match 50.0%; Score 51; DB 7; Length 83;  
 Best Local Similarity 100.0%; Pred. No. 1.2;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 RESLRNLRGY 20  
 |||||  
 Db 68 RESLRNLRGY 77

## RESULT 12

O9Y452 PRELIMINARY; PRT; 89 AA.  
 AC O9Y452;  
 DT 01-NOV-1999 (TREMBlrel. 12, Created)  
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
 DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)  
 DE HUMAN LEUCOCYTE ANTIGEN B (FRAGMENT).  
 GN HLA-B.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA CARTER V., DUNN P.P.;  
 FT "Identification of a novel HLA-B\*07 allele."  
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ237594; CAB40714.1; -.  
 FT NON\_TER 1 1  
 FT NON\_TER 89 89  
 SQ SEQUENCE 89 AA; 10609 MW; 659B735B CRC32;

Query Match 50.0%; Score 51; DB 4; Length 89;  
 Best Local Similarity 100.0%; Pred. No. 1.3;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 RESLRNLRGY 20  
 |||||  
 Db 74 RESLRNLRGY 83

## RESULT 13

P79615 PRELIMINARY; PRT; 89 AA.  
 AC P79615;  
 DT 01-MAY-1997 (TREMBlrel. 03, Created)  
 DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)  
 DE MHC CLASS I HLA-B (FRAGMENT).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA CHOPEK M., CAO K., ZHANG G.H.;  
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U88249; AAB48493.1; -.  
 DR PFAM; PF00129; MHC\_I; 1.  
 KW MHC.  
 FT NON\_TER 1 1  
 FT NON\_TER 89 89  
 SQ SEQUENCE 89 AA; 10489 MW; A7D3DF93 CRC32;

Query Match 50.0%; Score 51; DB 7; Length 89;  
 Best Local Similarity 100.0%; Pred. No. 1.3;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 RESLRNLRGY 20  
 |||||  
 Db 74 RESLRNLRGY 83

## RESULT 14

P79620 PRELIMINARY; PRT; 89 AA.  
 AC P79620;  
 DT 01-MAY-1997 (TREMBlrel. 03, Created)  
 DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)  
 DE MHC CLASS I HLA-B (FRAGMENT).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA CHOPEK M., CAO K., ZHANG G.H.;  
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U88254; AAB48498.1; -.  
 DR PFAM; PF00129; MHC\_I; 1.  
 KW MHC.  
 FT NON\_TER 1 1  
 FT NON\_TER 89 89  
 SQ SEQUENCE 89 AA; 10501 MW; FA3BFDCD CRC32;

Query Match 50.0%; Score 51; DB 7; Length 89;  
 Best Local Similarity 100.0%; Pred. No. 1.3;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 RESLRNLRGY 20  
 |||||  
 Db 74 RESLRNLRGY 83

## RESULT 15

O19548 PRELIMINARY; PRT; 89 AA.  
 AC O19548;  
 DT 01-JAN-1998 (TREMBlrel. 05, Created)  
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)  
 DE MHC CLASS I ANTIGEN (FRAGMENT).  
 GN HLA-B.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA CAO K., BURDETT L., ZHANG H., FERNANDEZ-VINA M.;  
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF016299; AAB69444.1; -.  
DR PFAM; PF00129; MHC\_I; 1.  
KW MHC.  
FT NON\_TER 1  
FT NON\_TER 89  
SQ SEQUENCE 89 AA; 10550 MW; 11452E40 CRC32;

Query Match 50.0%; Score 51; DB 7; Length 89;  
Best Local Similarity 100.0%; Pred. No. 1.3;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 RESLNLRGY 20  
| | | | | | | | | |  
Db 74 RESLNLRGY 83

Search completed: February 8, 2000, 19:16:13  
Job time: 21500 sec

**THIS PAGE BLANK (USPTO)**

OM of: US-08-653-294-27 to: GenEmbl.\* out\_format : pfs  
 Date: Feb 8, 2000 10:22 PM  
 About: Results were produced by the GenCore software, version 4.5,  
 Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:  
 -MODEL=frame+model+2n.model -DEV=xlp  
 -Q=/cgnl1/USPTO\_spool/US08653294/runat\_04022000\_160701\_15779/app\_query.fasta.2  
 -DB=GenEmbl -QFMT=fastcap -SUFFIX=rge -GAPOP=12.000 -GAPEXT=4.000  
 -MINNATCH=0.100 -LOOPEXT=0.000 -LOOPEXT=0.000 -GAPOP=4.500  
 -GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000  
 -KEYWORDS=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000  
 -DELEXT=7.000 -START=1 -MATRIX=blomsum62 -TRANS=human40.cdi  
 -LIST=45 -DOALIGN=200 -THR\_SCORE=pct -ALIGN=15 -MODE=LOCAL  
 -OUTFMT=pfs -NORM=ext -MINLEN=0 -MAXLEN=1000000 -USER=US08653294  
 -NCPU=6 -ICPU=3 -NO\_XLPXY -WAIT -THREADS=1

## Search information block:

Query: US-08-653-294-27

Query length: 20

Database: GenEmbl.\*

Database sequences: 821193

Database length: -1518192014

Search time (sec): 10176.920000

## score\_list:

Sequence	Strd Orig	ZScore	BScore	Len	Documentation
gb_p12:ATAC007134	+	52.00	85.99	84544	AC007134 Arabidopsis thaliana
gb_hlg6:AC010932	-	51.50	87.99	20373	AC010932 Homo sapiens chromo
gb_p12:HSHLABK1	+	51.00	137.28	243	U91332 Human cell line THAI DCH
gb_p12:HSHLABM1	+	51.00	137.28	243	U91332 Human cell line THAI DCH
gb_p12:HSHLABO1	+	51.00	137.28	243	U91336 Human cell line THAI DCH
gb_p12:AF014769	+	51.00	137.19	246	AF014769 Homo sapiens MHC class
gb_p12:AF014771	+	51.00	137.19	246	AF014771 Homo sapiens MHC class
gb_p12:AF014773	+	51.00	137.19	246	AF014773 Homo sapiens MHC class
gb_p12:AF014775	+	51.00	137.19	246	AF014775 Homo sapiens MHC class
gb_p12:AF014777	+	51.00	137.19	246	AF014777 Homo sapiens MHC class
gb_p12:AF014779	+	51.00	137.19	246	AF014779 Homo sapiens MHC class
gb_p12:AF014781	+	51.00	137.19	246	AF014781 Homo sapiens MHC class
gb_p12:AF014783	+	51.00	137.19	246	AF014783 Homo sapiens MHC class
gb_p12:AF014785	+	51.00	137.07	250	AF014785 Homo sapiens MHC class
gb_p12:AF014787	+	51.00	137.07	250	AF014787 Homo sapiens MHC class
gb_p12:AF014789	+	51.00	137.07	250	AF014789 Homo sapiens MHC class
gb_p12:AF015011	+	51.00	136.95	254	U59965 Human MHC class I proteol
gb_p12:HSB15011	+	51.00	136.49	270	D89333 Homo sapiens DNA for MHC
gb_p12:HSB15013	+	51.00	136.49	270	D89333 Homo sapiens DNA for MHC
gb_p12:HSB15015	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15017	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15019	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15021	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15023	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15025	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15027	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15029	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15031	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15033	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15035	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15037	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15039	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15041	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15043	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15045	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15047	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15049	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15051	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15053	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15055	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15057	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15059	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15061	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15063	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15065	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15067	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15069	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15071	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15073	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15075	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15077	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15079	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15081	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15083	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15085	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15087	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15089	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15091	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15093	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15095	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15097	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15099	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15101	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15103	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15105	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15107	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15109	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15111	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15113	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15115	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15117	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15119	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15121	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15123	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15125	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15127	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15129	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15131	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15133	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15135	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15137	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15139	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15141	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15143	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15145	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15147	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15149	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15151	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15153	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15155	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15157	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15159	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15161	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15163	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15165	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15167	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15169	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15171	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15173	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15175	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15177	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15179	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15181	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15183	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15185	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15187	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15189	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15191	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15193	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15195	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15197	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15199	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15201	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15203	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15205	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15207	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15209	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15211	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15213	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15215	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15217	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15219	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15221	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15223	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15225	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15227	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15229	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15231	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15233	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15235	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15237	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15239	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15241	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15243	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15245	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15247	+	51.00	136.49	270	AJ131852 Homo sapiens HLA-B gen
gb_p12:HSB15249	+</				

http://genome.wustl.edu/eddy/TrnScan-SE/. Simple repeats are identified by repeatmasker (Arian Smit, <http://ftp.genome.washington.edu/RW/RepeatMasker.html>). Regions of genomic sequence that are not annotated as genes but have predicted exons by GRAIL are annotated as misc features.

## FEATURES

```
source
1. 84544
/organism="Arabidopsis thaliana"
/cultivar="Columbia"
/db_xref="taxon:3702"
/chromosome="II"
/map="mi398"
/clone="F7H1"

misc_feature
1
/note="8353 nt before this point were not included in the submitted sequence, due to overlap with another BAC"
complement(1..4157)
/note="overlap with BAC clone F19G14 (AC006438:1..4157)."
928..1330
/gene="F7H1.1"
/note="similar to reverse transcriptase"
/pseudo
<2020..>5384
/gene="F7H1.2"
<2020..>6384
/gene="F7H1.2"
/note="F19G14.1"
2020..6384
/gene="F7H1.2"
/codon_start=1
/product="similar to retroviral aspartyl proteases"
/db_xref="GI:4678197"
/translation="MTVAVRTKSTRKAGTSVTRKSRSTGAVTTPNPPNRSR
ASRLTSSSGDTPQFFLHSADHFCNLNISRDLNTEINIGDSWAWLISLDKAKTIG
FIDGTLSPDLNRLMSRCKMVKSNLLNSPQIYRSILRMDASDIWRDLNRSR
FNTNPTNRYLQEIQDFRQGLTSLSEYTRKLTLDQDSTEALDEPCTCGKAMRL
QQAQAKIYKFLAGNESYAIRQIIIAKALPSLGEYHILDQNSQOSFNVAP
PAAFOVSEIPTVCYVQNGPKNGPICSFYNRVGHIAERCYKKGFPFGPT
KGAGKLOKPKLANVAESSEVNTSLEWGNLSKEOLOQFIAMFESQLONTPTT
YATASQSNLIGICSPSYISGILTVARHTLSSNTWIDSGATHVSHDRSLSS
LDTSVLNAIPGTVKISGVTGLKNDLILKNVIFPEFRNLNLSISLTDIGS
RVTFKNSCEIQDLIRGMIGORVANLYLDVDSISVNAVVDISMRHRLGHAS
LQRLDAISDLSLGRKNGSDFCHVCHAKORLSPFTSNKYCKEIFDLHIDVWGP
FSVTEVGYFTIYVDHSHRATWMLTKSEVLTFVPAFIOQVENVKVKVAVRS
DNPALKFTSYAEKGIYSPHSCPETPEQNSVVERKHQHLNVARALMFQSOVPLSLW
GCVLTVAFINRTPQLLNKTPYEILGTARVYEQLETFGLCYSSSPQRKHFO
PRRACLFLPGSGYKGLMDLESNTVFISRNQVHEEVFLPKNPGESSSLKLPYI
MVPVSGIISDTHSPSSLPQISDLPQISQVRKPPFAHLNDYHNTMQSDHKYPI
SSTISYKISPSHMCYINNTKIPTNTYAEQDTKEWCEAVDAEIGAMEKNTWETI
TLPKGVKAGCKWFTLKFLADNGLERYKARLVAKGYTOEGLDYDTTSPVAKMTTI
KLILKYASAKWFLKOLDVSNAFNGELEEEIFMKIPEGYAERKGLVLPNSVVLKLR
SIYGLQASQWKKRSSLLSGFKTHGDHFLFLKMYDGEFVIVYVDIVIAST
SEAAQLTELDORFKRLDGLKIFLGLVARTAGYSICORKYALELLELQSGMLA
CKVPVPMPLNKRKDDGLIEDIYQIRIVGKMLYITRPDIITFAVNLKQCFSSA
PRTHLTAAVRYQIKYIGVQGLFSASSDLTLKGFADSNWASQDSRRSTFTMF
VGDLSISWRKSKOHTVSRSAEAYRALATCEMWLFTLLVLSQASPPVTLYSDS
TRAITYATNPVFERHKKHLDCHTVRERLDNGLKLLHVRTEDQAVILTRPLFYQ
FEHLKMSILNIFSCSS"
10299..10369
/rpt_family="(GAAA)n"
join(<13346..13547,13635..>13735)
/gene="F7H1.3"
<13346..>13735
/gene="F7H1.3"
join(13346..13547,13635..13735)
/gene="F7H1.3"
/codon_start=1
/product="similar to replication protein A1"
/protein_id="A026944.1"
/db_xref="GI:4678198"
/translation="MVPGEWKATENGLIYAKFKATNHRKMGFMKATRVVRMEPLSD
SYLSTSTFDVLTGLNQNLIDVVGVQIVNVGEMETINVHNKPTKKINFLEURDHK"

repeat_region
14805..14919
/rpt_family="(CAT)n"
repeat_region
18429..18462
/rpt_family="(GAA)n"
<21092..>21322
/gene="F7H1.4"
<21092..>21322
/gene="F7H1.4"
/note="predicted by genscan and genefinder"
21092..21322
/gene="F7H1.4"
/note="hypothetical protein"
/codon_start=1
/protein_id="A026945.1"
/db_xref="GI:4678199"
/translation="MEALVYVHKVASEGGRPKTTVSDSGSHSRKSGKSFQVKDSPE
NRYATCKVDSLRSFSSSYNOERKQSEQR"
complement(21381..21407)
/rpt_family="POLY_A"
complement(<21932..>22627)
/gene="F7H1.5"
complement(<21932..>22627)
/gene="F7H1.5"
/note="predicted by genscan and genefinder"
complement(21932..22627)
/gene="F7H1.5"
/note="hypothetical protein"
/codon_start=1
/protein_id="A026946.1"
/db_xref="GI:4678200"
/translation="MERNVKKMLKRVISVFLSIGTVLMVIMLOTPKTCISPEAPSKP
HFFPRSTCDSPQROHLPLPKNARIWSSGANKSRISYSPFLRDLGFIQNHHTKA
LCISAGAGHMAALSOIGLSDVTAVELVDSIPLVKRADPHNLPEFPGVDFDEAFTALA
EALFPQVFEEMERTVRRGGFCVSVDECGDDVDIARLPHNSKVVVDVAVNTLESGK
KTSILFKVQDSPT"
join(<23753..24059,24489..>25330)
/gene="F7H1.6"
<23753..>25330
/gene="F7H1.6"
/note="predicted by genscan and genefinder"
join(23753..24059,24489..25330)
/gene="F7H1.6"
/note="hypothetical protein"
/codon_start=1
/protein_id="A026947.1"
/db_xref="GI:4678201"
/translation="MIGFNSDMKEHIREANKGEIHCHEPSSHKSONELTELLANDTKM
MLKKIKDKRYSVILDSIPDSVRKEQMTFLRCVDVSTCSKIBSEFFFLTHIKDKR
EYTDNPHGRSDVESLSETHGIGGFELFGIMLIWYDLAAVNIYSKISQFEDMDLEV
AISOGLGVLYLNKYTEGFEKAYESTQIAIEMKIAPVFKKSKKKKQFVEDVEKI
TSCMTAEESFRIDYFINIMDQAIMCIEIFQEVQYQIFGFLGVKRLKVAEDDELK
TSCMKLEASRLKHDVSDVDGELFMELKLDVLPKEITKPVLEVLKFLKIMDSYPTNT
WYAYRLTLTIPYVALAERTFSKLKILKYLKRLSTMSQERLNGALISV"
<28186..>28689
/gene="F7H1.7"
<28186..>28689
/gene="F7H1.7"
/note="predicted by genscan"
28186..28689
/gene="F7H1.7"
/note="hypothetical protein"

alignment_scores
Quality: 52.00 Length: 17
Ratio: 3.467 Gaps: 0
Percent Similarity: 88.235 Percent Identity: 52.941
alignment_block
US-08-653-294-27 x ATAC007134
Align seg 1/1 to: ATAC007134 from: 1 to: 84544
4 LeuAsnArgLeuSerGluArgArgGluSerLeuArgAsnLeuArgGlyTy 20
```

```

:::||||| |||:|||||:|||||:|||||:|||||:|||||:|||||
78760 AATAATAGATCAAGCCAGGCGTAACCAATTGAATAATCTTGGGTA 78809

```

20 r 20

78810 T 78810

seq\_name: gb\_htg6:AC010932

```

seq_documentation_block:
LOCUS AC010932 203273 bp DNA HTG 30-NOV-1999
DEFINITION Homo sapiens chromosome 15 clone RP11-296E22 map 15, *** SEQUENCING
IN PROGRESS ***, 36 unordered pieces.
ACCESSION AC010932
VERSION AC010932.2 GI:6479157
KEYWORDS HTG: HTGS_PHASE1.
SOURCE human.

```

## ORGANISM

```

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

## REFERENCE

```

1 (bases 1 to 203273)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,

```

```

Brown,A., Castle,A., Collangelo,M., Collins,S., Collymore,A.,
Cooke,P., DeLoreano,K., Dewar,K., Domino,M., Doneelan,L., Doyle,M.,

```

```

Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,

```

```

Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Lehoczky,J., Lien,C., Locke,K., Macdonald,P., Marquis,N.,

```

```

McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,

```

```

Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,

```

```

Testaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.

```

```

Direct Submission
Submitted (28-SEP-1999) Whitehead Institute/MIT Center for Genome

```

```

Research, 320 Charles Street, Cambridge, MA 02141, USA
On Nov 30, 1999 this sequence version replaced gi:5931431.

```

```

All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)

```

```

http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center

```

```

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR

```

```

Web site: http://www-seq.wi.mit.edu
Contact: sequence_submission@genome.wi.mit.edu

```

```

----- Project Information
Center project name: Lf684
Center clone name: 296_E-22

```

```

-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 36 contigs. The true order of the pieces

```

```

* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as

```

```

* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence

```

```

* as soon as it is available and the accession number will
* be preserved.

```

```

* 1 1375: contig of 1375 bp in length
* gap of unknown length

```

```

* 1376 3332: contig of 1957 bp in length
* gap of unknown length

```

```

* 3333 4414: contig of 1082 bp in length
* gap of unknown length

```

```

* 4415 6085: contig of 1671 bp in length
* gap of unknown length

```

```

* 6086 7636: contig of 1551 bp in length
* gap of unknown length

```

```

* 7637 9176: contig of 1540 bp in length
* gap of unknown length
* 9177 11451: contig of 2275 bp in length
* gap of unknown length
* 11452 14012: contig of 2561 bp in length
* gap of unknown length
* 14013 15805: contig of 1793 bp in length
* gap of unknown length
* 15806 17914: contig of 2109 bp in length
* gap of unknown length
* 17915 20644: contig of 2730 bp in length
* gap of unknown length
* 20645 23409: contig of 2765 bp in length
* gap of unknown length
* 23410 26283: contig of 2874 bp in length
* gap of unknown length
* 26284 28593: contig of 2310 bp in length
* gap of unknown length
* 28594 31564: contig of 2971 bp in length
* gap of unknown length
* 31565 34023: contig of 2459 bp in length
* gap of unknown length
* 34024 36890: contig of 2867 bp in length
* gap of unknown length
* 36891 41485: contig of 4595 bp in length
* gap of unknown length
* 41486 45667: contig of 4182 bp in length
* gap of unknown length
* 45668 49439: contig of 3771 bp in length
* gap of unknown length
* 49439 53273: contig of 3835 bp in length
* gap of unknown length
* 53274 56613: contig of 3340 bp in length
* gap of unknown length
* 56614 61902: contig of 5289 bp in length
* gap of unknown length
* 61903 67950: contig of 6048 bp in length
* gap of unknown length
* 67951 74487: contig of 6537 bp in length
* gap of unknown length
* 74488 82190: contig of 7703 bp in length
* gap of unknown length
* 82191 91448: contig of 9258 bp in length
* gap of unknown length
* 91449 98969: contig of 7521 bp in length
* gap of unknown length
* 98970 109518: contig of 10549 bp in length
* gap of unknown length
* 109519 119016: contig of 9498 bp in length
* gap of unknown length
* 119017 127923: contig of 8907 bp in length
* gap of unknown length
* 127924 139858: contig of 11935 bp in length
* gap of unknown length
* 139859 152524: contig of 12666 bp in length
* gap of unknown length
* 152525 164676: contig of 12152 bp in length
* gap of unknown length
* 164677 183255: contig of 18579 bp in length
* gap of unknown length
* 183256 203273: contig of 20018 bp in length.
Location/Qualifiers

```

```

FEATURES
Source

```

```

1. 203273
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="15"
/map="15"
/clone="RP11-296E22"
/clone_lib="RPC1-11 Human Male BAC"

```

```

BASE COUNT 54651 a 48192 c 46411 g 53915 t 104 others
ORIGIN

```

```

alignment_scores:
  Quality: 51.50      Length: 23
  Ratio: 2.861       Gaps: 1
  Percent Similarity: 78.261      Percent Identity: 52.174

alignment_block:
  US-08-653-294-27 x AC010932/rev ..
  Align seg 1/1 to reverse of: AC010932 from: 1 to: 203273
    1 TyrGlyArgLeuAsnArgLeuSerGluArgGlyTyr.....Le 14
      |||:||||| :|||:|||||:||||| ||| ||
  24858 TATGCAGGAATGGAGATCTTCAGATAAAGATATCTCTGTACCCCT 24809
    14 uArgAsnLeuArgGlyTyr 20
      |||:|||||:|||||:|||||
  24808 CAAAATCTCAGAGGCTAC 24790

seq_name: gb_pr2:HSHLABK1

seq_documentation_block:
  LOCUS HSHLABK1 243 bp DNA PRI 26-MAR-1997
  DEFINITION Human cell line THAI DCH022 MHC class I HLA-B gene (allele
  HLA-B*1521), exon 2.
  ACCESSION U91332
  VERSION U91332.1 GI:1906671
  KEYWORDS
  SEGMENT
  SOURCE
  ORGANISM Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
  Eutheria; Primates; Catarrhini; Hominidae; Homo.
  REFERENCE
  1 (bases 1 to 243)
  AUTHORS Chandanayingyong,D., Sirikong,M., Srinak,D., Longta,K.,
  Rungroung,E., Bejchandra,S., Juji,T., Tokunaga,K. and
  Grosse-Wilde,H.
  TITLE B15 alleles (B*1525)
  JOURNAL Unpublished
  REFERENCE
  2 (bases 1 to 243)
  AUTHORS Chandanayingyong,D., Sirikong,M., Srinak,D., Longta,K.,
  Rungroung,E., Bejchandra,S., Juji,T., Tokunaga,K. and
  Grosse-Wilde,H.
  TITLE Direct Submission
  JOURNAL Submitted (26-FEB-1997) Transfusion Medicine, Faculty of Medicine,
  Siriraj Hospital, Mahidol University, Prannok Road, Bangkok 10700,
  Thailand
  FEATURES
    source
      1..243
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /chromosome="6"
      /map="6p21"
      /cell_type="lymphoblastoid"
      /cell_line="THAI DCH022"
      1..243
      /gene="HLA-B"
      /note="similar to exon 2 of B*1501"
      /number=2
    BASE COUNT 52 a 79 c 84 g 28 t
    ORIGIN
      alignment_scores:
        Quality: 51.00      Length: 10
        Ratio: 5.100       Gaps: 0
        Percent Similarity: 100.000      Percent Identity: 100.000

      alignment_block:
        US-08-653-294-27 x HSHLABK1 ..
        Align seg 1/1 to: HSHLABK1 from: 1 to: 243
          11 ArgGluSerLeuArgAsnLeuArgGlyTyr 20
            |||:|||||:|||||:|||||:|||||
          195 CGAGAGAGCCTCGGAACCTCGCGGCTAC 224

seq_name: gb_pr2:HSHLABO1

seq_documentation_block:
  LOCUS HSHLABO1 243 bp DNA PRI 26-MAR-1997
  DEFINITION Human cell line THAI DCH012 MHC class I HLA-B gene (allele
  HLA-B*1525), exon 2.
  ACCESSION U91336
  VERSION U91336.1 GI:1906679
  KEYWORDS
  SEGMENT
  SOURCE
  ORGANISM Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
  Eutheria; Primates; Catarrhini; Hominidae; Homo.
  REFERENCE
  1 (bases 1 to 243)

```



AUTHORS Chandanayingyong,D., Sirikong,M., Srinak,D., Longta,K.,  
Rungroung,E., Bejchandra,S., Juji,T., Tokunaga,K. and  
Grosse-Wilde,H.  
TITLE B15 alleles (B\*1525)  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 243)  
AUTHORS Chandanayingyong,D., Sirikong,M., Srinak,D., Longta,K.,  
Rungroung,E., Bejchandra,S., Juji,T., Tokunaga,K. and  
Grosse-Wilde,H.  
TITLE Direct Submission  
JOURNAL Submitted (26-FEB-1997) Transfusion Medicine, Faculty of Medicine,  
Siriraj Hospital, Mahidol University, Prannok Road, Bangkok 10700,  
Thailand

FEATURES Location/Qualifiers  
source  
1..243  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="6"  
/map="6p21"  
/cell\_type="lymphoblastoid"  
1..243  
/gene="HLA-B"  
/note="Allele: HLA-B\*1525; similar to exon 2 of B\*1501"  
/number=2  
52 a 79 c 84 g 28 t

BASE COUNT  
ORIGIN

exon  
alignment\_scores  
Quality: 51.00 Length: 10  
Ratio: 5.100 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block  
US-08-653-294-27 x HSHLABO1 ..  
Align seg 1/1 to: HSHLABO1 from: 1 to: 243  
11 ArgGluSerLeuArgAsnLeuArgGlyTyr 20  
|||||  
195 CGAGAGAGCGCTGCGGAACCTGCGCGGCTAC 224

seq\_name: gb\_pr2:AF014769

seq\_documentation\_block: 246 bp DNA PRI 27-AUG-1997  
LOCUS AF014769  
DEFINITION Homo sapiens MHC class I antigen HLA-B gene (HLA-B\*1502 allele),  
exon 2 and partial cds.  
ACCESSION AF014769  
VERSION AF014769.1 GI:2345103  
KEYWORDS human.  
SOURCE  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 246)  
AUTHORS Chandanayingyong,D., Sirikong,M., Luangtrakool,K., Srinak,D.,  
Rungroung,E. and Bejchandra,S.  
TITLE B15 alleles (B\*1502)  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 246)  
AUTHORS Chandanayingyong,D., Sirikong,M., Luangtrakool,K., Srinak,D.,  
Rungroung,E. and Bejchandra,S.  
TITLE Direct Submission  
JOURNAL Submitted (18-JUL-1997) Transfusion Medicine, Faculty of Medicine,  
Siriraj Hospital, Mahidol University, Prannok Road, Bangkok 10700,  
Thailand

FEATURES Location/Qualifiers  
source  
1..246  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/cell\_line="Thai DCH4060"

gene  
/chromosome="6"  
/map="6p21"  
/cell\_type="lymphoblastoid"  
1..246  
/gene="HLA-B"  
/allele="HLA-B\*1502"  
1..246  
/gene="HLA-B"  
/note="2"  
/number=2  
1..246  
/gene="HLA-B"  
/codon\_start=3  
/product="MHC class I antigen HLA-B"  
/protein\_id="AAB67807.1"  
/db\_xref="GI:2345104"  
/translation="TAMSRGRGEPRPIAVGYVDDTQFVRFDSDAASPRMAPRAPWIE  
QEGPEYDRNTQISKNTQTYRESLRNLRGYNQSEA"  
BASE COUNT 55 a 81 c 82 g 28 t  
ORIGIN

alignment\_scores  
Quality: 51.00 Length: 10  
Ratio: 5.100 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block  
US-08-653-294-27 x AF014769 ..  
Align seg 1/1 to: AF014769 from: 1 to: 246  
11 ArgGluSerLeuArgAsnLeuArgGlyTyr 20  
|||||  
198 CGAGAGAGCGCTGCGGAACCTGCGCGGCTAC 227

seq\_name: gb\_pr2:AF014771

seq\_documentation\_block: 246 bp DNA PRI 27-AUG-1997  
LOCUS AF014771  
DEFINITION Homo sapiens MHC class I antigen HLA-B gene (HLA-B\*1502 allele),  
exon 2 and partial cds.  
ACCESSION AF014771  
VERSION AF014771.1 GI:2345107  
KEYWORDS human.  
SOURCE  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 246)  
AUTHORS Chandanayingyong,D., Sirikong,M., Luangtrakool,K., Srinak,D.,  
Rungroung,E. and Bejchandra,S.  
TITLE B15 alleles (B\*1502)  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 246)  
AUTHORS Chandanayingyong,D., Sirikong,M., Luangtrakool,K., Srinak,D.,  
Rungroung,E. and Bejchandra,S.  
TITLE Direct Submission  
JOURNAL Submitted (18-JUL-1997) Transfusion Medicine, Faculty of Medicine,  
Siriraj Hospital, Mahidol University, Prannok Road, Bangkok 10700,  
Thailand

FEATURES Location/Qualifiers  
source  
1..246  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/cell\_line="Thai DCH4061"  
/chromosome="6"  
/map="6p21"  
/cell\_type="lymphoblastoid"  
1..246  
/gene="HLA-B"  
/allele="HLA-B\*1502"  
1..246

gene  
exon

```

/feature="HLA-B"
/feature="2"
/number=2
<1..>246
/feature="HLA-B"
/codon_start=3
/product="MHC class I antigen HLA-B"
/protein_id="AAB67809.1"
/db_xref="GI:2345108"
/translation="TAMSPRGGEPRFIAGVYVDDTQFVRFDSDAASPRMAPRAPWIE
QEGPEYWDRTQISKNTQTYRESLRNLRGYNQSEA"
BASE COUNT      55 a      81 c      82 g      28 t
ORIGIN

```

```

alignment_scores:
  Quality: 51.00      Length: 10
  Ratio: 5.100      Gaps: 0
  Percent Similarity: 100.000      Percent Identity: 100.000

```

```

alignment_block:
US-08-653-294-27 x AF014771 ..

```

Align seg 1/1 to: AF014771 from: 1 to: 246

```

11 ArgGluSerLeuArgAsnLeuArgGlyTyr 20
|||||
198 CGAGAGAGCCTGCGGACCTGCGGGCTAC 227

```

seq\_name: gb\_pr2:AF014773

```

seq_documentation_block:
LOCUS AF014773      246 bp      DNA      PRI      27-AUG-1997
DEFINITION Homo sapiens MHC class I antigen HLA-B gene (HLA-B*1502 allele),
exon 2 and partial cds.

```

```

ACCESSION AF014773
VERSION AF014773.1 GI:2345111
KEYWORDS
SOURCE human.

```

```

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 246)
AUTHORS Chandanayingyong,D., Sirikong,M., Luangtrakool,K., Srinak,D.,
Rungroung,E. and Bejchandra,S.
TITLE B15 alleles (B*1502)
JOURNAL Unpublished

```

```

REFERENCE 2 (bases 1 to 246)
AUTHORS Chandanayingyong,D., Sirikong,M., Luangtrakool,K., Srinak,D.,
Rungroung,E. and Bejchandra,S.
TITLE Direct Submission
JOURNAL Submitted (18-JUL-1997) Transfusion Medicine, Faculty of Medicine,
Siriraj Hospital, Mahidol University, Prannok Road, Bangkok 10700,
Thailand

```

```

FEATURES
Source Location/Qualifiers
1..246
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="Thai DCH3086"
/chromosome="6"
/map="6p21"
/cell_type="lymphoblastoid"
<1..>246
/gene="HLA-B"
/allele="HLA-B*1502"

```

```

1..246
/feature="HLA-B"
/feature="2"
/number=2
<1..>246
/gene="HLA-B"
/codon_start=3
/product="MHC class I antigen HLA-B"

```

```

/feature="HLA-B"
/feature="2"
/number=2
<1..>246
/gene="HLA-B"

```

```

/feature="HLA-B"
/feature="2"
/number=2
<1..>246
/gene="HLA-B"

```

```

/feature="HLA-B"
/feature="2"
/number=2
<1..>246
/gene="HLA-B"

```

```

/protein_id="AAB67811.1"
/db_xref="GI:2345112"
/translation="TAMSPRGGEPRFIAGVYVDDTQFVRFDSDAASPRMAPRAPWIE
QEGPEYWDRTQISKNTQTYRESLRNLRGYNQSEA"
BASE COUNT      55 a      81 c      82 g      28 t
ORIGIN

```

```

alignment_scores:
  Quality: 51.00      Length: 10
  Ratio: 5.100      Gaps: 0
  Percent Similarity: 100.000      Percent Identity: 100.000

```

```

alignment_block:
US-08-653-294-27 x AF014773 ..

```

Align seg 1/1 to: AF014773 from: 1 to: 246

```

11 ArgGluSerLeuArgAsnLeuArgGlyTyr 20
|||||
198 CGAGAGAGCCTGCGGACCTGCGGGCTAC 227

```

seq\_name: gb\_pr2:AF014775

```

seq_documentation_block:
LOCUS AF014775      246 bp      DNA      PRI      27-AUG-1997
DEFINITION Homo sapiens MHC class I antigen HLA-B gene (HLA-B*1502 allele),
exon 2 and partial cds.

```

```

ACCESSION AF014775
VERSION AF014775.1 GI:2345115
KEYWORDS
SOURCE human.

```

```

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 246)
AUTHORS Chandanayingyong,D., Sirikong,M., Luangtrakool,K., Srinak,D.,
Rungroung,E. and Bejchandra,S.
TITLE B15 alleles (B*1502)
JOURNAL Unpublished

```

```

REFERENCE 2 (bases 1 to 246)
AUTHORS Chandanayingyong,D., Sirikong,M., Luangtrakool,K., Srinak,D.,
Rungroung,E. and Bejchandra,S.
TITLE Direct Submission
JOURNAL Submitted (18-JUL-1997) Transfusion Medicine, Faculty of Medicine,
Siriraj Hospital, Mahidol University, Prannok Road, Bangkok 10700,
Thailand

```

```

FEATURES
source Location/Qualifiers
1..246
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="Thai DCH018"
/chromosome="6"
/map="6p21"
/cell_type="lymphoblastoid"
<1..>246
/gene="HLA-B"
/allele="HLA-B*1502"

```

```

1..246
/feature="HLA-B"
/feature="2"
/number=2
<1..>246
/gene="HLA-B"

```

```

1..246
/feature="HLA-B"
/feature="2"
/number=2
<1..>246
/gene="HLA-B"

```

```

1..246
/feature="HLA-B"
/feature="2"
/number=2
<1..>246
/gene="HLA-B"

```

```

1..246
/feature="HLA-B"
/feature="2"
/number=2
<1..>246
/gene="HLA-B"

```

```

1..246
/feature="HLA-B"
/feature="2"
/number=2
<1..>246
/gene="HLA-B"

```

```

1..246
/feature="HLA-B"
/feature="2"
/number=2
<1..>246
/gene="HLA-B"

```

```

1..246
/feature="HLA-B"
/feature="2"
/number=2
<1..>246
/gene="HLA-B"

```

alignment\_scores:  
 Quality: 51.00 Length: 10  
 Ratio: 5.100 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-08-653-294-27 x AF014775 ..

Align seg 1/1 to: AF014775 from: 1 to: 246

11 ArgGluSerLeuArgAsnLeuArgGlyTyr 20  
 |||||  
 198 CGAGAGCCTGCGGACCTGCGGCTAC 227

seq\_name: gb\_pr2:AF014777

seq\_documentation\_block: 246 bp DNA PRI 27-AUG-1997  
 LOCUS AF014777 Homo sapiens MHC class I antigen HLA-B gene (HLA-B\*1502 allele),  
 DEFINITION exon 2 and partial cds.

ACCESSION AF014777.1 GI:2345119

## KEYWORDS

SOURCE human.

## ORGANISM

Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 246)

AUTHORS Chandanayingyong,D., Sirikong,M., Luangtrakool,K., Srinak,D.,

Rungroung,E. and Bejchandra,S.

TITLE B15 alleles (B\*1502)

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 246)

AUTHORS Chandanayingyong,D., Sirikong,M., Luangtrakool,K., Srinak,D.,

Rungroung,E. and Bejchandra,S.

TITLE Direct Submission

JOURNAL Submitted (18-JUL-1997) Transfusion Medicine, Faculty of Medicine,

Siriraj Hospital, Mahidol University, Prannok Road, Bangkok 10700,

Thailand

## FEATURES

source Location/Qualifiers  
 1..246  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /cell\_line="Thai DCH017"  
 /chromosome="6"  
 /map="6p21"  
 /cell\_type="lymphoblastoid"  
 <1..>246

## gene

/gene="HLA-B"  
 /allele="HLA-B\*1502"

## exon

/gene="HLA-B"  
 /note="2"

## CDS

<1..>246  
 /gene="HLA-B"  
 /codon\_start=3  
 /product="MHC class I antigen HLA-B"  
 /protein\_id="AAB67815.1"  
 /db\_xref="GI:2345120"  
 /translation="TAMSRPGRGEPRFIAVGVDQTFVRFDSDAASPRMAPRAPWIE  
 QEGPEYDNTQISKNTQTYRESLNLRGYNQSEA"

BASE COUNT 55 a 81 c 82 g 28 t

## ORIGIN

alignment\_scores:  
 Quality: 51.00 Length: 10  
 Ratio: 5.100 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-08-653-294-27 x AF014777 ..

Align seg 1/1 to: AF014777 from: 1 to: 246

11 ArgGluSerLeuArgAsnLeuArgGlyTyr 20  
 |||||  
 198 CGAGAGCCTGCGGACCTGCGGCTAC 227

seq\_name: gb\_pr2:AF014779

seq\_documentation\_block: 246 bp DNA PRI 27-AUG-1997

LOCUS AF014779 Homo sapiens MHC class I antigen HLA-B gene (HLA-B\*1502 allele),  
 DEFINITION exon 2 and partial cds.

ACCESSION AF014779.1 GI:2345123

## KEYWORDS

SOURCE human.

## ORGANISM

Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 246)

AUTHORS Chandanayingyong,D., Sirikong,M., Luangtrakool,K., Srinak,D.,

Rungroung,E. and Bejchandra,S.

TITLE B15 alleles (B\*1502)

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 246)

AUTHORS Chandanayingyong,D., Sirikong,M., Luangtrakool,K., Srinak,D.,

Rungroung,E. and Bejchandra,S.

TITLE Direct Submission

JOURNAL Submitted (18-JUL-1997) Transfusion Medicine, Faculty of Medicine,

Siriraj Hospital, Mahidol University, Prannok Road, Bangkok 10700,

Thailand

## FEATURES

source Location/Qualifiers  
 1..246  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /cell\_line="Thai DCH002"  
 /chromosome="6"  
 /map="6p21"  
 /cell\_type="lymphoblastoid"  
 <1..>246

## gene

/gene="HLA-B"  
 /allele="HLA-B\*1502"

## exon

/gene="HLA-B"  
 /note="2"

## CDS

<1..>246  
 /gene="HLA-B"  
 /codon\_start=3  
 /product="MHC class I antigen HLA-B"  
 /protein\_id="AAB67817.1"  
 /db\_xref="GI:2345124"  
 /translation="TAMSRPGRGEPRFIAVGVDQTFVRFDSDAASPRMAPRAPWIE  
 QEGPEYDNTQISKNTQTYRESLNLRGYNQSEA"

BASE COUNT 55 a 81 c 82 g 28 t

## ORIGIN

alignment\_scores:  
 Quality: 51.00 Length: 10  
 Ratio: 5.100 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-08-653-294-27 x AF014779 ..

Align seg 1/1 to: AF014779 from: 1 to: 246

11 ArgGluSerLeuArgAsnLeuArgGlyTyr 20  
 |||||  
 198 CGAGAGCCTGCGGACCTGCGGCTAC 227

```

seq_name: gb_pr2:AF014781
seq_documentation_block:
LOCUS AF014781 246 bp DNA PRI 27-AUG-1997
DEFINITION Homo sapiens MHC class I antigen HLA-B gene (HLA-B*1502 allele),
exon 2 and partial cds.
ACCESSION AF014781
VERSION AF014781.1 GI:2345127
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 246)
Chandanayingyong,D., Sirikong,M., Luangtrakool,K., Srinak,D.,
Rungroung,E. and Bejchandra,S.
B15 alleles (B*1502)
TITLE
JOURNAL
REFERENCE
2 (bases 1 to 246)
Chandanayingyong,D., Sirikong,M., Luangtrakool,K., Srinak,D.,
Rungroung,E. and Bejchandra,S.
Direct Submission
AUTHORS
Submitted (18-JUL-1997) Transfusion Medicine, Faculty of Medicine,
Siriraj Hospital, Mahidol University, Prannok Road, Bangkok 10700,
Thailand
FEATURES
Location/Qualifiers
source
1..246
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="Thai DCH003"
/chromosome="6"
/map="6p21"
/cell_type="lymphoblastoid"
<1..>246
/gene="HLA-B"
/allele="HLA-B*1502"
1..246
/gene="HLA-B"
/feature="2"
/number=2
<1..>246
/gene="HLA-B"
/codon_start=3
/product="MHC class I antigen HLA-B"
/protein_id="AAB67819.1"
/db_xref="GI:2345128"
/translation="TAMSRPGRGPRFTIAGVYDDTQVRFDSDAASPRMAPRWIE
QEGPEYDWRNTQISKNTQTYRSLRLRGYNQSEA"
BASE COUNT 55 a 81 c 82 g 28 t
ORIGIN

alignment_scores:
Quality: 51.00 Length: 10
Ratio: 5.100 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-27 x AF014781 ..
Align seg 1/1 to: AF014781 from: 1 to: 246
11 ArgGluserLeuArgAsnLeuArgGlyTyr 20
|||||
198 CGAGAGAGCCTGCGGACCTGCGGCGCTAC 227

seq_name: gb_pr2:AF014783
seq_documentation_block:
LOCUS AF014783 250 bp DNA PRI 27-AUG-1997
DEFINITION Homo sapiens MHC class I antigen HLA-B gene (HLA-B*1525 allele),
exon 2 and partial cds.
ACCESSION AF014783
VERSION AF014783.1 GI:2345135
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

REFERENCE 1 (bases 1 to 250)  
 AUTHORS Chandanayingyong,D., Sirikong,M., Luangtrakool,K., Srinak,D.,  
 TITLE Rungroung,E. and Bejchandra,S.  
 JOURNAL B15 alleles (B\*1525)  
 REFERENCE 2 (bases 1 to 250)  
 AUTHORS Chandanayingyong,D., Sirikong,M., Luangtrakool,K., Srinak,D.,  
 TITLE Rungroung,E. and Bejchandra,S.  
 JOURNAL Direct Submission  
 TITLE Submitted (18-JUL-1997) Transfusion Medicine, Faculty of Medicine,  
 JOURNAL Siriraj Hospital, Mahidol University, Prannok Road, Bangkok 10700,  
 Thailand

FEATURES  
 source Location/Qualifiers  
 1..250

/organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /cell\_line="Thai DCH1109"  
 /chromosome="6"  
 /map="6p21"  
 /cell\_type="lymphoblastoid"

1..250  
 /gene="HLA-B"  
 /allele="HLA-B\*1525"  
 1..250  
 /gene="HLA-B"  
 /note="2"

/number=2  
 <1..>250  
 /gene="HLA-B"

/codon\_start=1  
 /product="MHC class I antigen HLA-B"  
 /protein\_id="AAB67823.1"

/db\_xref="GI:2345136"  
 /translation="FYTAMSRPGRGEPFIAVGVDVDTQFVRFSDAASPRMAPRAPW  
 IEQEGPEYWDRETOISKNTQTYRESLNRLGYNQSEA"

BASE COUNT 54 a 81 c 84 g 31 t  
 ORIGIN

alignment\_scores:  
 Quality: 51.00 Length: 10  
 Ratio: 5.100 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-08-653-294-27 x AF014785

Align seg 1/1 to: AF014785 from: 1 to: 250

11 ArgGluSerLeuArgAsnLeuArgGlyTyr 20  
 |||||  
 202 CGAGAGAGCCTGCGGAACCTGCGGCGCTAC 231

seq\_name: gb\_pr2:AF014787

seq\_documentation\_block:  
 LOCUS AF014787 250 bp DNA PRI 27-AUG-1997  
 DEFINITION Homo sapiens MHC class I antigen HLA-B gene (HLA-B\*1525 allele),  
 exon 2 and partial cds.  
 ACCESSION AF014787  
 VERSION AF014787.1 GI:2345139  
 KEYWORDS  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 250)  
 AUTHORS Chandanayingyong,D., Sirikong,M., Luangtrakool,K., Srinak,D.,  
 TITLE Rungroung,E. and Bejchandra,S.  
 JOURNAL B15 alleles (B\*1525)  
 REFERENCE 2 (bases 1 to 250)  
 AUTHORS Chandanayingyong,D., Sirikong,M., Luangtrakool,K., Srinak,D.,

TITLE  
 JOURNAL

Rungroung,E. and Bejchandra,S.

Direct Submission  
 Submitted (18-JUL-1997) Transfusion Medicine, Faculty of Medicine,  
 Siriraj Hospital, Mahidol University, Prannok Road, Bangkok 10700,  
 Thailand

FEATURES  
 source Location/Qualifiers  
 1..250

/organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /cell\_line="Thai DCH3258"  
 /chromosome="6"  
 /map="6p21"  
 /cell\_type="lymphoblastoid"

1..250  
 /gene="HLA-B"  
 /allele="HLA-B\*1525"  
 1..250  
 /gene="HLA-B"  
 /note="2"

/number=2  
 <1..>250  
 /gene="HLA-B"

/codon\_start=1  
 /product="MHC class I antigen HLA-B"  
 /protein\_id="AAB67825.1"

/db\_xref="GI:2345140"  
 /translation="FYTAMSRPGRGEPFIAVGVDVDTQFVRFSDAASPRMAPRAPW  
 IEQEGPEYWDRETOISKNTQTYRESLNRLGYNQSEA"

BASE COUNT 54 a 81 c 84 g 31 t  
 ORIGIN

alignment\_scores:

Quality: 51.00 Length: 10  
 Ratio: 5.100 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-08-653-294-27 x AF014787

Align seg 1/1 to: AF014787 from: 1 to: 250

11 ArgGluSerLeuArgAsnLeuArgGlyTyr 20  
 |||||  
 202 CGAGAGAGCCTGCGGAACCTGCGGCGCTAC 231

**THIS PAGE BLANK (USPTO)**

OM of: US-08-653-294-27 to: N\_Geneseq\_36:\* out\_format : pfs  
Date: Feb 8, 2000 7:31 PM  
About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:  
-MODEL=framet.p2n.model -DEV=xlp  
-O=/cgnl\_1/USPTO\_spool/US08653294/runat\_04022000\_160701\_15807/app\_query.fasta.2  
-DB=N\_Geneseq\_36 -QFMT=fastap -SUFFIX=ring -GAPOP=12.000  
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000  
-GAPOP=4.500 -GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500  
-GAPOP=6.000 -GAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500  
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR\_SCORE=pct  
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0  
-MAXLEN=1000000 -USER=US08653294 -NCPU=6 -ICPU=3 -NO\_XLPXY -WAIT  
-THREADS=1

## Search information block:

Query: US-08-653-294-27  
Query length: 20  
Database: N\_Geneseq\_36:\*  
Database sequences: 311585  
Database length: 125096042  
Search time (sec): 873.190000

## score\_list:

Sequence	Strid	Orig	ZScore	Escore	Len	Documentation
N_Geneseq_36:Q12115	-	51.00	137.59	4.60	1089	HLA-B35 exon ..
N_Geneseq_36:Q75974	-	51.00	124.88	23.49	4059	HLA-B7 expression vector. New
N_Geneseq_36:Q75973	-	51.00	122.94	30.15	4965	PHLA-B7/beta-2 microglobulin e
N_Geneseq_36:Q22494	+	48.00	119.62	46.11	2675	Human CRP1TO-related gene CR-3
N_Geneseq_36:Q71145	+	48.00	119.62	46.11	2675	DNA encoding human CRP1TO-rela
N_Geneseq_36:Q70987	+	48.00	119.62	46.11	2675	CRP1TO-related gene, CR-3. Ass
N_Geneseq_36:Q783924	+	48.00	119.62	46.11	2675	Human CRP1TO-related gene (CR-
N_Geneseq_36:Q22495	+	48.00	119.22	119.26	5761	Human CRP1TO gene CR-1. New CR
N_Geneseq_36:Q83923	+	48.00	112.22	119.26	5761	Human CRP1TO-related gene (CR-
N_Geneseq_36:Q17562	+	47.50	121.81	34.85	1818	Coding sequence for the alpha
N_Geneseq_36:Q12977	+	47.00	123.16	29.29	1346	Enterococcus faecalis genome c
N_Geneseq_36:Q89291	+	46.00	134.93	6.48	289	EST clone CG336. New polynucleo
N_Geneseq_36:Q82783	+	46.00	125.78	20.94	745	Clone dn809.5 isolated from hum
N_Geneseq_36:Q99594	+	46.00	117.00	64.56	1849	Mouse TIE-2 receptor ligand 3
N_Geneseq_36:Q11549	+	46.00	97.74	763.38	13585	Tumour rejection antigen prec
N_Geneseq_36:Q08510	+	45.00	112.31	117.87	2182	Sequence of amylase gene and u
N_Geneseq_36:Q17563	+	44.50	111.99	122.69	1920	Coding sequence for the alpha
N_Geneseq_36:Q90017	+	44.00	137.55	4.63	116	EST clone CE242. New polynucleo
N_Geneseq_36:Q12116	+	44.00	115.82	75.14	1101	HLA-C exon Cb-1. HLA-C gene, D
N_Geneseq_36:Q12117	+	44.00	115.82	75.14	1101	HLA-C exon Cb-2. HLA-C gene, D
N_Geneseq_36:Q48111	-	44.00	110.81	142.80	1849	Nucleotide sequence encoding C
N_Geneseq_36:Q19798	-	44.00	108.52	191.57	2344	Human growth regulator protein
N_Geneseq_36:Q69921	+	44.00	103.20	378.99	4066	Expression vector pNF2232. New
N_Geneseq_36:Q52426	+	44.00	102.44	417.68	4398	Streptococcus pneumoniae gene
N_Geneseq_36:Q69919	+	44.00	100.90	509.34	5162	Expression vector pNF2176. New
N_Geneseq_36:Q69922	+	44.00	100.90	509.34	5162	Mutagenic plasmid pNF2237. New
N_Geneseq_36:Q69927	+	44.00	100.71	521.59	5262	Promoter probe vector pNF2283.
N_Geneseq_36:Q69920	+	44.00	98.16	723.11	6850	Expression vector pNF2214. New
N_Geneseq_36:Q85924	+	44.00	96.61	882.21	8043	Human protein tyrosine phospho
N_Geneseq_36:Q20556	+	44.00	93.65	1.3e+03	10928	Polynucleotide sequence from
N_Geneseq_36:Q20835	+	43.00	111.15	136.76	1296	Synthetic glycosylase oxidoredu
N_Geneseq_36:Q20833	+	43.00	110.96	140.04	1321	Manipulated glycosylase oxidore
N_Geneseq_36:Q20834	+	43.00	110.96	140.04	1321	Modified glycosylase oxidoreduc
N_Geneseq_36:Q78564	+	43.00	108.93	181.82	1631	CPRI-GOX gene fusion. Chemical
N_Geneseq_36:Q20832	+	43.00	108.59	189.86	1689	Glycosylase oxidoreductase gene
N_Geneseq_36:Q22705	+	43.00	108.57	190.28	1692	Glycosylase oxidoreductase gene
N_Geneseq_36:Q52137	-	43.00	104.00	241.95	2716	Streptococcus pneumoniae genom
N_Geneseq_36:Q83943	+	43.00	89.81	2.1e+03	11811	Bacterial artificial chromoso
N_Geneseq_36:Q12984	-	43.00	86.24	3.3e+03	17087	Enterococcus faecalis genome
N_Geneseq_36:Q20248_06	+	43.00	68.26	3.3e+04	110000	Continuation (7 of 10) of
N_Geneseq_36:Q20250	+	43.00	68.14	3.3e+04	111309	Borrelia burgdorferi polynud

N\_Geneseq\_36:X41493 - 42.50 120.57 40.82 416 ! Human secreted protein 5' ES  
N\_Geneseq\_36:N50226 + 42.50 81.18 6.4e+03 24593 ! Sequence of opine synthase  
N\_Geneseq\_36:N50182 + 42.50 81.18 6.4e+03 24596 ! Complete nucleotide sequen  
N\_Geneseq\_36:X07410 + 42.00 106.84 237.76 1470 ! Saccharothrix australiensis

seq\_name: N\_Geneseq\_36:Q12115

seq\_documentation\_block:

ID Q12115 standard; DNA; 1089 BP.

AC Q12115;

DT 29-AUG-1991 (first entry)

DE HLA-B35 exon.

KW Human leukocyte antigen; probe; major histocompatibility complex;

MHC; class I; ss.

OS Homo sapiens.

PH Key Location/Qualifiers

FT cds 1..1089

FT /\*tag= a

FN J03112486-A.

PD 14-MAY-1991.

PE 22-SEP-1989; 247697

PR 22-SEP-1989; JP-247697.

PA (Olyu ) OLYMPUS OPTICAL KK.

DR WPI; 91-182991/25.

DR P-PSDB; R12464.

PT HLA-B35 gene - used in DNA probe and transformant cells for

immunising animals, for developing monoclonal antibody.

PS Claim 1; Page 1; lipp; Japanese.

CC Probes comprising part of the sequence can be used to identify

Class I genes. The DNA can be expressed for immunisation of

animals and prodn. of monoclonal antibodies specific for the

HLA-B35 antigen. See also J03112485 and J03112487.

SQ Sequence . 1089 BP; 221 A; 336 C; 359 G; 173 T;

alignment\_scores:

Quality: 51.00 Length: 10

Ratio: 5.100 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-08-653-294-27 x Q12115 ..

Align seg 1/1 to: Q12115 from: 1 to: 1089

11 ArgGluSerLeuArgAsnLeuArgGlyTyr 20

295 CGAGAGAGCGCTCGGAACCTGCGCGCTAC 324

seq\_name: N\_Geneseq\_36:Q75974

seq\_documentation\_block:

ID Q75974 standard; cDNA; 4059 BP.

AC Q75974;

DT 23-AUG-1995 (first entry)

DE pHLA-B7 expression vector.

KW expression vector; pHLA-B7; heavy human HLA-B7; bicistronic mRNA;

light beta-2 microglobulin; class I major histocompatibility complex;

MHC; human leukocyte antigen; HLA; covalently closed circular DNA; ds.

OS Synthetic.

PH Key Location/Qualifiers

FT misc\_feature 1..354

FT /\*tag= a

/note= "pBR322 backbone contg. bacterial origin of

replication"

FT cds 355..1170

FT /\*tag= b

/note= "kanamycin resistance gene open reading frame;

the gene is taken from the transposable element

FT polyA\_signal tn303

FT complement (1410..1177)

FT /\*tag= c

/note= "SV40 polyA signal sequence"

```

FT intron complement (1412..1560)
FT /tag= d
FT /note= "SV40 small t intron"
FT 3'utr complement (1561..1794)
FT /tag= e
FT /note= "3' untranslated region of HLA-B7
FT heavy chain mRNA"
FT cds complement (1795..2880)
FT /tag= f
FT /note= "HLA-B7 open reading frame"
FT ltr complement (2886..3415)
FT /tag= g
FT /note= "Rous sarcoma virus 3' LTR promoter region"
FT misc_feature 3416..4059
FT /tag= h
FT /note= "pBR322 backbone"
FT W09429469-A.
PD 22-DEC-1994.
PF 27-MAY-1994; U06069.
PR 07-JUN-1993; US-074344.
PA (UNMI ) UNIV MICHIGAN.
PA (VICA-) VICAL INC.
PI Lew D, Marquet M, Nabel EG, Nabel GJ;
DR WPI; 95-036494/05.
PT New vectors for gene therapy, partic for tumours - comprising
PT genetic material encoding one or more cistron(s) which express
PT immunogenic or therapeutic peptide(s)
PS Claim 9; Page 42-43; 50pp; English.
CC This HLA-B7 antigen encoding plasmid was developed to incorporate many
CC advantageous features, eg. the kanamycin resistance gene. The
CC eradication of two open reading frames encoding portions of SV40 viral
CC proteins lowers the risk of tumorigenicity. The vector may also operate
CC as a cassette into which cistrons may be inserted and removed at will
CC for the transcription and subsequent translation of peptides of interest.
CC The vector is used partic. for the treatment of neoplastic disease,
CC eg. melanoma, and provides enhanced gene delivery and expression
CC in vivo.
SQ Sequence 4059 BP; 975 A; 1051 C; 1033 G; 1000 T;

alignment_scores:
Quality: 51.00 Length: 10
Ratio: 5.100 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-27 x Q75974/rev ..
Align seg 1/1 to reverse of: Q75974 from: 1 to: 4059
11 ArgGluSerLeuArgAsnLeuArgGlyTyr 20
|||||
2589 CGAGAGAGCCCTGCGGACCTGCGGCTAC 2560

seq_name: N_Geneseq_36:Q75973
seq_documentation_block:
ID Q75973 standard; CDNA; 4965 BP.
AC Q75973;
DT 23-AUG-1995 (first entry)
DE pHLA-B7/beta-2 microglobulin expression vector.
KW expression vector; pHLA-B7/beta-2 microglobulin; heavy human HLA-B7;
KW light beta-2 microglobulin; class I major histocompatibility complex;
KW MHC; bicistronic mRNA; human leukocyte antigen; HLA;
KW covalently closed circular DNA; ds.
OS Synthetic.
FH Key Location/Qualifiers
FT ltr 1..529
FT /tag= a
FT /note= "Rous sarcoma virus LTR promoter domain, derived
FT for the Schmidt-Rupin strain nucleotides
FT 8673-9146. This region also includes a 56 bp
FT region of a synthetic oligonucleotide which

```

```

FT modifies this regulatory sequence to effect a
FT higher level of expression of downstream
FT sequences. The oligonucleotide removes a
FT polyadenylation signal sequence originally found
FT in the RSV DNA sequence."
FT misc_signal 531..534
FT /tag= b
FT /label= consensus_Kozak_signal_sequence
FT 535..1620
FT /tag= c
FT /note= "HLA-B7 heavy chain open reading frame"
FT 535..606
FT /tag= d
FT /note= "encodes putative signal peptide of the
FT HLA-B7 heavy chain"
FT 607..1620
FT /tag= e
FT /note= "encodes putative HLA-B7 heavy chain mature
FT peptide"
FT 1621..1853
FT /tag= f
FT /note= "3' untranslated sequence of HLA-B7 heavy
FT chain mRNA"
FT 1854..1888
FT /tag= g
FT /note= "Multiple cloning site, forms a junction
FT between the HLA-B7 sequence and the EMCV-CITE
FT sequence, and is used to facilitate subcloning"
FT 1889..2479
FT /tag= h
FT /note= "murine encephalomyocarditis CAP-independent
FT translational enhancer (EMCV-CITE); taken from
FT nucleotides 255-843 of cloned EMCV genomic DNA.
FT It is a non-coding regulatory sequence functioning
FT as an internal entry point for the eukaryotic
FT ribosomal subunits when located within a mRNA
FT the beta-2 microglobulin, downstream of the HLA-B7
FT stop codon on this bicistronic mRNA to be
FT recognised by the ribosome"
FT 2480..2839
FT /tag= i
FT /note= "encodes beta-2 microglobulin; this cDNA is
FT deriv. from chimpanzee (differs to the human
FT cDNA by only 4 bases)"
FT 2840..2846
FT /tag= j
FT /note= "3' untranslated region of the beta-2
FT microglobulin mRNA"
FT 2847..2870
FT /tag= k
FT /note= "synthetic linker"
FT 3'UTR ..3111
FT /tag= l
FT /note= "bovine growth hormone 3'UTR and transcriptional
FT terminator; it starts at a blunt-ended BglI site
FT within the 3'UTR of the mRNA coding sequence"
FT 2979..2984
FT /tag= m
FT 3112..3151
FT /tag= n
FT /note= "synthetic linker to facilitate cloning"
FT complement (3151..3967)
FT /tag= o
FT /note= "kanamycin resistance gene open reading frame;
FT the gene is taken from the transposable element
FT Tn903"
FT 4014..4965
FT /tag= p
FT /note= "pBR322 backbone contg. bacterial origin of
FT replication, it represents nucleotides 2244-3193"
FT W09429469-A.
PD 22-DEC-1994.

```



```

PF 27-MAY-1994; U06069.
PR 07-JUN-1993; US-074344.
PA (UNMI ) UNIV MICHIGAN.
PA (VICA-) VICAL INC.
PI Lew D., Marquet M., Nabel EG, Nabel GJ;
DR WPI; 95-036494/05.
PT New vectors for gene therapy, partic for tumours - comprising
PT genetic material encoding one or more cistron(s) which express
PT immunogenic or therapeutic peptide(s)
PS Claim 8; Page 41-42; 50pp; English.
CC The pHA-B7/beta-2 microglobulin plasmid expression vector, in addition
CC to the kanamycin resistance gene, contains the plasmid DNA encoding the
CC heavy (HLA-B7) and light (beta-2 microglobulin) proteins of a
CC class I major histocompatibility complex (MHC) antigen. The plasmid is
CC designed to express these two proteins via a bicistronic mRNA in
CC eukaryotic cells. Initiation of transcription of the mRNA is dependent
CC on a Kous Sarcoma Virus (RSV) promoter sequence deriv. from the 3' long
CC terminal repeat. Termination of transcription is dependent upon the
CC polyadenylation signal sequence deriv. from the bovine growth hormone
CC gene. Eukaryotic cell translation of the heavy chain is regulated by the
CC 5' cap-dependent protein start site. Translation of the light chain is
CC controlled by the CITE. Finally the replication of the plasmid in
CC bacterial cells is controlled by the presence of a bacterial origin of
CC replication. The vector is used partic. for the treatment of neoplastic
CC disease, eg. melanoma, and provides enhanced gene delivery and expression
CC in vivo.
SQ Sequence 4965 BP; 1171 A; 1293 C; 1338 G; 1163 T;

alignment_scores:
  Quality: 51.00 Length: 10
  Ratio: 5.100 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-27 x Q75973 ..
Align seg 1/1 to: Q75973 from: 1 to: 4965

11 ArgGluSerLeuArgAsnLeuArgGlyTyr 20
|||||
829 CGAGAGAGCTGCGGAACCTGCGCGCTAC 858

seq_name: N_Geneseq_36:Q22494

seq_documentation_block:
ID Q22494 standard; DNA; 2675 BP.
AC Q22494;
DT 28-AUG-1992 (first entry)
DE Human CRIPTO-related gene CR-3.
KW cell proliferation; tumour; CR-1; transforming growth factor;
KW epidermal growth factor; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT cds 809..1375
FT /*tag= a
FT /product= CR-3
PN US7749001-A.
PD 25-FEB-1992.
PF 23-AUG-1991; 749001.
PR 23-AUG-1991; US-749001.
PA (USSH ) US DEPT HEALTH & HUMAN.
PI Salomon D, Persico M;
DR WPI; 92-123675/15.
DR P-PSDB: R22547.
PT New CRIPTO gene CR-1 and CRIPTO-related gene CR-3 genomic DNA -
PT CR-3 protein and anti-CR-3 antibodies, useful in immunoassay to
PT detect CR-3 as tumour specific marker
PS Disclosure; Page 29; 44pp; English.
CC The CR-3 nucleotide sequence was isolated from a human genomic
CC library using a labelled DNA fragment containing 800bp upstream of
CC the translation initiation site of CR-1. Clones were isolated which
CC contained the 5' cDNA non-coding region of CR-3. One clone was

```

```

CC sequenced and found to include a complete CRIPTO cDNA lacking
CC introns and containing a poly(A) tract at the 3' end. Seven single
CC base pair substitutions were observed in the coding region (C.f.
CC CR-1 sequence in Ciccodicola A. et al., EMBO J.8:1987-1991(1989)),
CC six of which gave rise to amino acid changes. The 3' non-coding
CC sequence is less similar (97% identical) to CR-1. Most of the
CC differences occur in the inverted Alu sequence. The unusual poly(A)
CC addition site ACTAAA found in the CR-1 gene is also conserved in
CC CR-3. The similarity between CR-1 and CR-3 extends for 697
CC nucleotides upstream of the initiator AUG where it is possible to
CC observe 7 base pair substitutions and 6 nucleotide deletions.
CC The 14 "Others" in the sequence represent nucleotides whose
CC identity, although determined by the inventors, was unclear in
CC the sequence printed in the specification.
SQ Sequence 2675 BP; 738 A; 611 C; 553 G; 759 T;

alignment_scores:
  Quality: 48.00 Length: 21
  Ratio: 3.200 Gaps: 1
Percent Similarity: 71.429 Percent Identity: 57.143

alignment_block:
US-08-653-294-27 x Q22494 ..
Align seg 1/1 to: Q22494 from: 1 to: 2675

2 GlyArgLeuAsnArgSerGlu.....ArgArgGluSerLeuArgAs 16
|||||
2161 GCGAGACTGGGTAGGAAGAGGAGCAATAAGAGAGAGAGGTTGAAAAA 2210

16 nLeuArgGlyTyr 20
|||||
2211 CAAATGGGTTAC 2223

seq_name: N_Geneseq_36:T71145

seq_documentation_block:
ID T71145 standard; DNA; 2675 BP.
AC T71145;
DT 19-AUG-1997 (first entry)
DE DNA encoding human CRIPTO-related gene product, CR-3.
KW CRIPTO; CR1; CR3; epidermal growth factor superfamily; EGF; mitogen;
KW tumour marker; epithelium; mesenchyme; diagnosis; prognosis; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT cds 809..1375
FT /*tag= a
FT /product= CR-3
PN US5620866-A.
PD 15-APR-1997.
PF 23-AUG-1991; 749001.
PR 23-AUG-1991; US-749001.
PR 17-NOV-1993; US-154198.
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PI Persico MG, Salomon DS;
DR WPI; 97-235180/21.
DR P-PSDB: W19980.
PT New DNA encoding the human CRIPTO-related gene product CR-3 - used
PT to produce recombinant protein, useful for preparation of diagnostic
PT antibodies
PS Claim 1; Column 17-20; 25pp; English.
CC T71145 encodes a human CRIPTO-related gene product CR-3. CR-3 is
CC thought to be a tumour-specific marker and as such DNA and amino
CC acid sequences of CR-3 and derivatives can be used in the
CC diagnosis, prognosis and possibly treatment of some types of cancer.
CC CR-3 includes a region homologous to products of the epidermal
CC growth factor superfamily of genes and it may be a mitogen involved
CC in regulating proliferation, differentiation and transformation of
CC mesenchymal and epithelial cells.
SQ Sequence 2675 BP; 739 A; 618 C; 555 G; 763 T;

```





```

CC Modification of the seed storage protein profile can result in the
CC production of novel soy protein products with unique and valuable
CC functional characteristics.
SQ Sequence 1818 BP; 581 A; 424 C; 427 G; 386 T;

alignment_scores:
  Quality: 47.50 Length: 23
  Ratio: 2.794 Gaps: 1
  Percent Similarity: 73.913 Percent Identity: 47.826

alignment_block:
US-08-653-294-27 x V17562 ..

Align seg 1/1 to: V17562 from: 1 to: 1818

1 TyrGLYArg.....LeuAsnArgLeuSerGluArgGluSerLe 14
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
631 TATGGTGCATTCGGCGCTCCAGAGGGTTCAACCAACGCTCCCACT 680

14 uArgAsnLeuArgGlyTyr 20
|:::||||||| |||
681 TCAGAAATCTCGAGACTTAC 699

seq_name: N_Geneseq_36:X12977

seq_documentation_block:
ID X12977 standard; DNA; 1346 BP.
AC X12977;
DT 19-MAR-1999 (first entry)
DE Enterococcus faecalis genome contig SEQ ID NO:40.
KW Enterococcus faecalis; contig; detection; Enterococcal infection;
KW vaccine; attenuation; computer readable medium; ds.
OS Enterococcus faecalis.
PN WC9850555-A2.
PD 12-NOV-1998.
PF 04-MAY-1998; U08985.
PR 14-NOV-1997; US-066009.
PR 06-MAY-1997; US-044031.
PR 16-MAY-1997; US-046655.
PA (HUMA-) HUMAN GENOME SCI INC.
PI Barash SC, Dillon PJ, Kunsch CA;
PT WPI; 99-04317/704.
PT New isolated Enterococcus faecalis polynucleotides and polypeptides
PT - used to develop products for the detection of Enterococcus and for
PT use in vaccines for prevention or attenuation of Enterococcus
PT infection.
PS Claim 1; Page 412-413; 2084pp; English.
CC A computer readable medium has been developed which has recorded on it
CC 982 nucleotide sequences isolated from the Enterococcus faecalis genome.
CC X12938 to X13919 represent these nucleotide sequences which are primary
CC nucleotide sequences, also known as contigs. The computer-based system
CC can identify fragments of the Enterococcus faecalis genome with
CC commercial importance. The products can be used to detect the presence
CC of Enterococcus faecalis in samples. They can also be used for
CC diagnosing Enterococcal infection in an animal and monitoring
CC progression of disease, and for identifying agents which can be used to
CC modulate the growth or pathogenicity of Enterococcus faecalis, or
CC another related organism, in vivo or in vitro. In particular the
CC polypeptides encoded by the Enterococcus faecalis nucleotide sequences
CC can be used in vaccines to prevent or attenuate an Enterococcal
CC infection.
SQ Sequence 1346 BP; 395 A; 269 C; 284 G; 395 T;

```

Align seg 1/1 to reverse of: X12977 from: 1 to: 1346

1 TvrGlyArgLeuAsnArgLeuSerGluArgGluSerLeuArgAsnLe 17  
 |||||  
 1202 TATGGAGATTAGAAAGTTCGTCATTAAAGACAGTATCGAAATT 1153  
 |||||  
 17 uArgGly 19  
 :  
 1152 TAAAGC 1146

seq\_name: N\_Geneseq\_36:V89291

seq\_documentation\_block:

ID V89291 standard; cDNA: 289 BP.  
 AC V89291;  
 DE EST clone CG336.  
 KW Human; secreted protein; expressed sequence tag; EST; haematopoiesis;  
 KW tissue growth; activin; inhibin; chemotaxis; chemokinesis; haemostatic;  
 KW receptor; ligand; thrombolytic; anti-inflammatory; cadherin; anti-tumour;  
 KW gene therapy; ss.  
 OS Homo sapiens.  
 PN W09845436-A2.  
 PD 15-OCT-1998.  
 PF 10-APR-1998; U06955.  
 PR 10-APR-1997; US-838821.  
 PA (GEM) GENETICS INST INC.  
 PI Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D,  
 PI Racie LA, Spaulding V, Treacy M;  
 DR WPI: 99-070077/06.  
 PT New polynucleotides encoding human secreted proteins - derived from  
 PT e.g. human blood, kidney, foetal lung, placenta, testes, brain,  
 PT ovary, pituitary, retina and colon cDNA libraries.  
 PS Claim 1: Page 171; 618pp; English.  
 CC The present sequence represents a human expressed sequence tag (EST).  
 CC The polynucleotide, which is a secreted EST, and the encoded protein  
 CC are predicted to have useful biological activities which would make  
 CC them suitable for treating, preventing or ameliorating medical  
 CC conditions in humans and animals, although no supporting data is  
 CC given. Suggested activities include nutritional activity, immune  
 CC stimulating or suppressing activity, haematopoiesis regulating  
 CC activity, tissue growth activity, activin/inhibin activity,  
 CC chemotactic/chemokinetic activity, haemostatic and thrombolytic  
 CC activity, receptor/ligand activity, anti-inflammatory activity,  
 CC cadherin/tumour invasion suppressor activity, tumour inhibition  
 CC activity. The polynucleotide may also be useful for gene therapy.  
 SQ Sequence 289 BP; 63 A; 63 C; 63 G; 100 T;

alignment\_scores:

Quality: 46.00 Length: 13  
 Ratio: 4.182 Gaps: 0  
 Percent Similarity: 84.615 Percent Identity: 76.923

alignment\_block:

US-08-653-294-27 x V89291 ..

Align seg 1/1 to: V89291 from: 1 to: 289

3 ArgLeuAsnArgLeuSerGluArgGluSerLeuArg 15  
 |||||  
 74 AGAGTCAACAGGCTAAGTGAGAGGAGGAGGATGCTACGC 112

seq\_name: N\_Geneseq\_36:V82783

seq\_documentation\_block:

ID V82783 standard; cDNA: 745 BP.  
 AC V82783;  
 DE 25-FEB-1999 (first entry)  
 DE Clone dn809\_5 isolated from human foetal brain cDNA library.  
 KW Secreted protein; nutritional activity; immune stimulating; vaccine;  
 KW suppressing activity; haematopoiesis regulating activity;  
 KW tissue growth activity; activin; inhibin activity; chemotactaxis;

KW chemokinetic activity; haemostasis; thrombolytic activity; receptor;  
 KW ligand; anti-inflammatory; cadherin; tumour invasion suppressor;  
 KW tumour inhibition; gene therapy; ds.  
 OS Homo sapiens.  
 PN W09842739-A2.  
 PD 01-OCT-1998.  
 PF 20-MAR-1998; U05653.  
 PR 19-MAR-1998; US-044466.  
 PR 21-MAR-1997; US-822167.  
 PA (GEM) GENETICS INST INC.  
 PI Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D,  
 PI Racie LA, Spaulding V, Treacy M;  
 DR WPI: 98-609890/51.  
 DR P-PSDB: W85461.  
 PT New polynucleotides encoding secreted human proteins - derived from  
 PT human foetal brain, adult brain, foetal kidney, placenta or adult  
 PT pineal gland cDNA libraries.  
 PS Claim 26; Page 80; 113pp; English.  
 CC The present sequence encodes a secreted protein. The polynucleotide and  
 CC secreted protein are predicted to have biological activities which would  
 CC make them suitable for treating, preventing or ameliorating medical  
 CC conditions in humans and animals, although no supporting data is given.  
 CC Suggested activities include nutritional activity, immune stimulating  
 CC (e.g. as vaccines) or suppressing activity, haematopoiesis regulating  
 CC activity, tissue growth activity, activin/inhibin activity,  
 CC chemotactic/chemokinetic activity, haemostatic and thrombolytic activity,  
 CC receptor/ligand activity, anti-inflammatory activity, cadherin/tumour  
 CC invasion suppressor activity, and tumour inhibition activity (no data is  
 CC given in the specification to support these activities). The  
 CC polynucleotide is also stated to be useful for gene therapy.  
 SQ Sequence 745 BP; 179 A; 229 C; 176 G; 161 T;

alignment\_scores:

Quality: 46.00 Length: 13  
 Ratio: 4.182 Gaps: 0  
 Percent Similarity: 84.615 Percent Identity: 76.923

alignment\_block:

US-08-653-294-27 x V82783/rev ..

Align seg 1/1 to reverse of: V82783 from: 1 to: 745

3 ArgLeuAsnArgLeuSerGluArgGluSerLeuArg 15  
 |||||  
 123 AGAGTCAACAGGCTAAGTGAGAGGAGGAGGATGCTACGC 85

seq\_name: N\_Geneseq\_36:T99594

seq\_documentation\_block:

ID T99594 standard; cDNA: 1849 BP.  
 AC T99594;  
 DE 22-JUN-1998 (first entry)  
 DE Mouse TIE-2 receptor ligand 3 cDNA.  
 KW TIE-2 receptor ligand; TIE ligand-3; tyrosine kinase; mouse;  
 KW angiogenesis; vascularisation; blood vessel growth; tumour;  
 KW therapy; receptorbody; ss.  
 OS Mus musculus.  
 FH Key Location/Qualifiers  
 FT CDS 47..1576  
 FT /\*tag= a

PN W09748804-A2.  
 PD 24-DEC-1997.  
 PF 19-JUN-1997; U10728  
 PR 02-AUG-1996; US-022999.  
 PR 19-JUN-1996; US-665926.  
 PR 02-JUL-1996; US-021087.  
 PA (REGG-) REGENERON PHARM INC.  
 PI Jones PF, Valenzuela DM, Yancopoulos GD;  
 DR WPI: 98-063143/06.  
 DR P-PSDB: W26792.  
 PT DNA encoding TIE ligands 3 and 4 - useful for, e.g. blocking blood  
 PT vessel growth and promoting neovascularisation, etc.

CC The tumour rejection antigen precursor (TRAP) is processed to a  
CC tumour rejection antigen (TRA) by HLA-A2 molecules. TRA  
CC is used to generate cytotoxic T lymphocytes for treating cancer (esp.  
CC melanoma). It can also be used to raise specific antibodies, and  
CC when complexed with HLA-A2, it can be used to produce vaccines.  
CC Cytotoxic T lymphocytes so generated can be used in adoptive  
CC transfer or generated, or they can be generated *in vivo* by using a  
CC vector containing the appropriate gene or using TRA or TRAP together  
CC with an adjuvant that facilitates entry into HLA-A2 presenting cells  
CC Diagnostic methods involving the detection of expression of TRAP  
CC can be used in the detection of cancers.  
SQ Sequence 13585 bp; 3827 A; 2859 C; 2968 G; 3894 T;

```
alignment_scores:      Length: 16
                       Quality: 46.00
                       Ratio: 3.538
                       Gaps: 0
Percent Similarity: 81.250
Percent Identity: 56.250
```

```
alignment_block:
US-08-653-294-27 x T11549
..
```

4 LeuAsnArgLeuSerGluArgGluSerLeuArgAsnLeuArgGly 19  
:::|||||::||| ::||| |||::: ||| ||| |||  
246 ATCAACAGAAATTCTCCGCAACGTTCAGTCTCCAACCTCAGAGGG 293

3 ArgLeuAsnArgLeuSerGluArgGluSerLeuArgAsnLeuArgG1 19  
:::||||| |||:::|||||:::|||||:::|||||:::|||||  
707 CAGCTTAACAGCCTCCAAGCAGAGAGGGGAACAACCTGCACAGTCTCCTGGG 756

19 yTyr 20  
| :::  
.

757 CCAT 760

```
seq_documentation_block:
ID T11549 standard; DNA; 13585 BP.
```

AC T11549;  
ID T11549 SCANDIAG, DNA, 199603 BF.

DT 18-APR-1996 (first entry)

DE Tumour rejection antigen precursor coding sequence.

KW Tumour rejection antigen precursor; TRAP; TRA; melanoma; cancer;  
KW Tumour rejection antigen precursor; TRAP; TRA; melanoma; cancer;

KW tumour; treatment; detection; vaccine; HLA-A2; adoptive transfer;  
KW tumour; treatment; detection; vaccine; HLA-A2; adoptive transfer;

KW T cell; T lymphocyte; human leukocyte antigen; ds.

OS Homo sapiens.

FH	Key	Location/Qualifiers
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	6	6
7	7	7
8	8	8
9	9	9
10	10	10
11	11	11
12	12	12
13	13	13
14	14	14
15	15	15
16	16	16
17	17	17
18	18	18
19	19	19
20	20	20
21	21	21
22	22	22
23	23	23
24	24	24
25	25	25
26	26	26
27	27	27
28	28	28
29	29	29
30	30	30
31	31	31
32	32	32
33	33	33
34	34	34
35	35	35
36	36	36
37	37	37
38	38	38
39	39	39
40	40	40
41	41	41
42	42	42
43	43	43
44	44	44
45	45	45
46	46	46
47	47	47
48	48	48
49	49	49
50	50	50
51	51	51
52	52	52
53	53	53
54	54	54
55	55	55
56	56	56
57	57	57
58	58	58
59	59	59
60	60	60
61	61	61
62	62	62
63	63	63
64	64	64
65	65	65
66	66	66
67	67	67
68	68	68
69	69	69
70	70	70
71	71	71
72	72	72
73	73	73
74	74	74
75	75	75
76	76	76
77	77	77
78	78	78
79	79	79
80	80	80
81	81	81
82	82	82
83	83	83
84	84	84
85	85	85
86	86	86
87	87	87
88	88	88
89	89	89
90	90	90
91	91	91
92	92	92
93	93	93
94	94	94
95	95	95
96	96	96
97	97	97
98	98	98
99	99	99
100	100	100

FT	misc_feature	2685
----	--------------	------

```
FT=
/*tag= a
```

FT	note= "Unidentified nucleotide."
0.00	0.00
0.01	0.01
0.02	0.02
0.03	0.03
0.04	0.04
0.05	0.05
0.06	0.06
0.07	0.07
0.08	0.08
0.09	0.09
0.10	0.10
0.11	0.11
0.12	0.12
0.13	0.13
0.14	0.14
0.15	0.15
0.16	0.16
0.17	0.17
0.18	0.18
0.19	0.19
0.20	0.20
0.21	0.21
0.22	0.22
0.23	0.23
0.24	0.24
0.25	0.25
0.26	0.26
0.27	0.27
0.28	0.28
0.29	0.29
0.30	0.30
0.31	0.31
0.32	0.32
0.33	0.33
0.34	0.34
0.35	0.35
0.36	0.36
0.37	0.37
0.38	0.38
0.39	0.39
0.40	0.40
0.41	0.41
0.42	0.42
0.43	0.43
0.44	0.44
0.45	0.45
0.46	0.46
0.47	0.47
0.48	0.48
0.49	0.49
0.50	0.50
0.51	0.51
0.52	0.52
0.53	0.53
0.54	0.54
0.55	0.55
0.56	0.56
0.57	0.57
0.58	0.58
0.59	0.59
0.60	0.60
0.61	0.61
0.62	0.62
0.63	0.63
0.64	0.64
0.65	0.65
0.66	0.66
0.67	0.67
0.68	0.68
0.69	0.69
0.70	0.70
0.71	0.71
0.72	0.72
0.73	0.73
0.74	0.74
0.75	0.75
0.76	0.76
0.77	0.77
0.78	0.78
0.79	0.79
0.80	0.80
0.81	0.81
0.82	0.82
0.83	0.83
0.84	0.84
0.85	0.85
0.86	0.86
0.87	0.87
0.88	0.88
0.89	0.89
0.90	0.90
0.91	0.91
0.92	0.92
0.93	0.93
0.94	0.94
0.95	0.95
0.96	0.96
0.97	0.97
0.98	0.98
0.99	0.99
1.00	1.00

FT	misc_feature	9422. .9456
FT		

/\*tag= b

/note= "This region is 4.7-5.3 kilobases in

but the sequence of this region has

ET	810	deducted.
ET	11530	

misc_feature	11539	C
ET	/44302	C
ET		

```

File
  / *tag= c
  / notes "identified nucleotide "

```

FT PN WO9601557-A1 /note= "Unidentified nucleotide."

PN WO9801337-A1.  
PD 25-JAN-1996.

23 JAN 1990.  
27 JUN 1995: U08153.

PR 08-JUL-1994: US-272351.

PR 10-JAN-1995; US-370319.

PA (LUDW-) LUDWIG INST CANCER RES.

PI Boon-falleur T, Brichard V, Coulie P, De PLAEN I

PI Traversari C, Van PEL A, Wolfel T;

DR WPI; 96-097390/10.

PT Genomic DNA encoding a tumour rejection antigen precursor -

processed to antigen presented by HLA-A2, useful for treating or

PT	diagnosing melanoma	17	100
23	100	100	100

PS Claim 3; Page 18-22; 4lpp; English.

OM of: US-08-653-294-27 to: EST:\* out\_format : pfs

Date: Feb 8, 2000 6:23 AM

About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODE=frameat\_p2n.model -DEV=xlp  
-O=/cgn1\_1/USPTO.spool/US08653294/runat\_04022000\_160700\_15770/app\_query.fasta.2  
-DB=EST -QFM=fastap -SUFFIX=rst -GAPOP=12.000 -GAPEXT=4.000  
-MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -QGAPOP=4.500  
-GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000  
-FGAEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000  
-DEEXT=7.000 -START=1 -MATRIX=blosum62 -TRANS-human40.cdi  
-LIST=45 -DCCALIGN=200 -THR SCORE=pct -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pfs -NORM=ext -MINLEN=0 -MAXLEN=1000000 -USER=US08653294  
-NCPU=6 -ICPU=3 -NO\_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-08-653-294-27

Query length: 20

Database: EST.\*

Database sequences: 4538634

Database length: 1887831982

Search time (sec): 7600.090000

score\_list:

Sequence	Strd Orig	ZScore	Escore	Len	Documentation
gb_est12:AA319533	+	51.00	6.30	103	AA319533 EST21772 Adrenal gland
gb_est13:AA361477	+	51.00	135.06	189	AA361477 EST70761 T-cell lymph
gb_est13:AA352603	+	51.00	13.84	103	AA352603 EST60621 Activated T-c
gb_est12:AA294911	+	51.00	148.49	218	AA352603 EST60621 Activated T-c
gb_est13:AA352960	+	51.00	20.81	259	AA294911 EST100074 Pancreas tum
gb_est13:AA352960	+	51.00	146.33	270	AA352960 EST611017 Activated T-c
gb_est8:AA058454	+	51.00	33.38	373	AA058454 z167407.r1 Stratagene
gb_est11:AA224068	+	51.00	142.43	397	AA224068 zrl11f02.r1 Stratagene
gb_est9:AA160317	+	51.00	142.23	405	AA160317 z056c07.r1 Stratagene
gb_est10:AA100680	+	51.00	141.89	419	AA100680 z190502.r1 Stratagene
gb_est7:WA0489	+	51.00	141.70	427	WA0489 zc84b01.r1 Pancreatic is
gb_est12:AA310808	+	51.00	141.33	443	AA310808 EST161593 Jurkat T-cell
gb_est9:CI8310	+	51.00	140.62	475	CI8310 C18310 Human placenta ch
gb_est17:AA663896	+	51.00	140.54	479	AA663896 ae74d01.s1 Stratagene
gb_est35:AL039796	+	51.00	140.04	503	AL039796 DKFZ434B1912.r1 434
gb_est28:AI524732	+	51.00	138.55	583	AI524732 th12c03.x1 NCI_CGAP_C
gb_est37:AW008453	+	51.00	137.30	560	AW008453 wv55d05.x1 NCI_CGAP_C
gb_est11:AA263135	+	51.00	136.56	710	AA263135 PM05598 KGI-a Lambda Z
gb_gss3:B20346	-	51.00	135.31	804	B20346 F18J1-Sp6 IGF Arabidopsi
gb_gss10:AA248102	-	50.50	137.61	553	AA248102 HS_2015.A1_H09 MR CIT
gb_gss8:AA009853	+	50.00	139.60	523	AA009853 HS_2173.A1_G06 MF CIT
gb_gss3:B50728	-	50.00	136.69	523	B50728 CIT-HSP-63C12.IV CIT-HS
gb_gss13:AA0475632	+	50.00	134.64	641	AA0475632 CITBI-EI-2590L18.TF CT
gb_gss13:AA0475632	+	50.00	133.50	717	AA0475632 CITBI-EI-2590L16.TF CT
gb_gss12:AA0377514	-	49.50	137.00	438	AA0377514 RPII-11-162P10.TJ RPII
gb_est5:DI73898	-	49.00	140.37	471	DI73898 CELK064C5F Yuji Kohara
gb_est9:AA063950	+	49.00	136.57	395	AA063950 m143f02.r1 Stratagene
gb_gss11:AA029878	+	49.00	134.31	494	AA029878 HS_3005.B2.A11.T7 CIT
gb_est11:AA205710	-	49.00	133.77	521	AA205710 zq6a12.s1 Stratagene
gb_est34:AA195124	+	48.50	142.17	37.46	AA195124 EST7504.Y1 Gm-c1010 G
gb_est31:AA331579	+	48.00	138.77	57.91	AA331579 EST35512 Embryo, 8 wee
gb_est44:AW184423	-	48.00	132.50	441	AW184423 fj14a03.y1 Zebrafish a
gb_est35:AT827894	-	48.00	132.05	461	AT827894 wf12b01.x1 Soares NFL
gb_est10:AA129456	+	48.00	131.56	484	AA129456 zn78c06.s1 Stratagene
gb_gss15:AA0641570	+	48.00	131.07	508	AA0641570 RCR933-EcorI-319.TP RP
gb_gss15:AA0641570	+	48.00	130.83	520	AA0641570 RCR933-EcorI-319.TP RP
gb_est11:AA216757	-	48.00	130.72	162.65	AA216757 zq96b01.s1 Stratagene
gb_est44:AW187318	-	48.00	130.27	172.33	AW187318 ENLGH112771 Six-day Cd
gb_est15:AA526987	-	48.00	130.16	174.77	AA526987 n119b06.s1 NCI_CGAP_C
gb_gss11:AA0295324	+	48.00	129.78	183.36	AA0295324 HS_3111.A1_E09 MR CIT
gb_gss15:AA0641860	+	48.00	128.63	212.68	AA0641860 RCR933-EcorI-6121.TV H
gb_est34:AI794913	+	47.50	131.30	151.00	AI794913 sb73d11.y1 Gm-c1010 G

gb\_est34:AI794877 + 47.50 129.22 197.12 527 ! AI794877 sb72h01.y1 Gm-c1010  
gb\_est34:AI795011 + 47.50 129.22 197.12 527 ! AI795011 sb75a10.y1 Gm-c1010  
gb\_gss6:AO873097 + 47.50 128.80 207.84 549 ! AO873097 V59C8 mTn-3xHA/lacZ

seq\_name: gb\_est12:AA319533

seq\_documentation\_block: 103 bp mRNA EST 19-APR-1997  
LOCUS AA319533  
DEFINITION Adrenal gland tumor Homo sapiens cDNA 5' end similar to  
Similar to major histocompatibility complex, class I, B  
(GB:W6102), mRNA sequence.

ACCESSION AA319533  
VERSION AA319533  
KEYWORDS EST.  
SOURCE human.

ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 103)

AUTHORS

Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Feldner,R.A.,  
Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,  
White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,  
Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,  
Fitzgerald,L.M., Fritchman,J.L., Geoghagen,N.S.,  
Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,E.S., Jr.,  
Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,  
Moreno-Palacios,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,  
Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,  
Small,K.V., Spriggs,T.A., Uterback,T.R., Weidman,J.F., Li,Y.,  
Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,  
Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,E.J.,  
He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,  
Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Weissner,P.S., Olsen,H.,  
Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M.,  
Dillon,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,  
Fraser,C.M. and Venter,J.C.

TITLE

Initial assessment of human gene diversity and expression patterns  
based upon 83 million nucleotides of cDNA sequence  
Nature 377 (6547 Suppl), 3-174 (1995)

MEDLINE

12140200

COMMENT On Sep 12, 1996 this sequence version replaced gi:1393672.

Other\_ESTs: THC169519

Contact: Kerlavage, AR

Bioinformatics

The Institute for Genomic Research

9712 Medical Center Drive, Rockville, MD 20850 USA

Tel: 3018699056

Fax: 3018699423

Email: arkerlavet@tigr.org

For clone availability, additional sequence and expression

information related to this EST, please check the TIGR Human Gene

Index (<http://www.tigr.org/tdb/hgi/hgi.html>)

Seq primer: M13 Reverse.

Location/Qualifiers

source

1. .103

/organism="Homo sapiens"

/db\_xref="ATCC (inhost):120083"

/db\_xref="taxon:9606"

/clone\_lib="Adrenal gland tumor"

/dev\_stage="adult"

/note="Organ: adrenal gland; Vector: pBluescript SK-;

Site:1: EcorI; Site:2: XhoI"

BASE COUNT 26 a 35 c 29 g 13 t

ORIGIN

alignment\_scores:

Quality: 51.00 Length: 10

Ratio: 5.100 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-08-653-294-27 x AA319533

Align seg 1/1 to: AA319533 from: 1 to: 103

```

11 ArgGlusSerLeuArgAsnLeuArgGlyTyr 20
|||||
27 CGAGAGAGCGCTCGGACCTTCGCGCTAC 56

```

seq\_name: gb\_est13:AA361477

seq\_documentation\_block: 189 bp mRNA EST 21-APR-1997  
 LOCUS AA361477  
 DEFINITION EST70761 T-cell lymphoma Homo sapiens cDNA 5' end similar to similar to major histocompatibility complex, class I, Bw62.3, mRNA sequence.

ACCESSION AA361477  
 VERSION AA361477.1 GI:2013795  
 KEYWORDS EST.

SOURCE

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 189)

AUTHORS Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A., Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C., Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D., Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S., Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr., Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M., Moreno-Palauques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M., Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R., Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y., Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J., Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A., He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K., Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Meissner,P.S., Olsen,H., Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M., Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and Venter,J.C.

TITLE Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence

JOURNAL Nature 377 (6547 Suppl), 3-174 (1995)

MEDLINE 12140200

COMMENT On Sep 12, 1996 this sequence version replaced gi:1404737.

Other\_ESTs: THCI69519

Contact: Kerlavage, AR

Bioinformatics

The Institute for Genomic Research

9712 Medical Center Drive, Rockville, MD 20850 USA

Tel: 3018699056

Fax: 3018699423

Email: arkerlavet@tigr.org

For clone availability, additional sequence and expression

information related to this EST, please check the TIGR Human Gene

Index (<http://www.tigr.org/tdb/hgi/hgi.html>)

Seq primer: M13 Reverse.

FEATURES

Location/Qualifiers

1..189  
 /organism="Homo sapiens"  
 /db\_xref="ATCC (inhost):165623"  
 /db\_xref="taxon:9606"  
 /clone\_lib="T-cell lymphoma"  
 /cell\_type="T-lymphocyte"  
 /note="Vector: pBluescript SK-; Site\_1: EcoRI; Site\_2: XhoI"

BASE COUNT 45 a 60 c 55 g 25 t 4 others

ORIGIN

alignment\_scores:

Quality: 51.00 Length: 10

Ratio: 5.100 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-08-653-294-27 x AA361477

Align seg 1/1 to: AA361477 from: 1 to: 189

```

11 ArgGlusSerLeuArgAsnLeuArgGlyTyr 20
|||||
37 CGAGAGAGCGCTCGGACCTTCGCGGTAC 66

```

seq\_name: gb\_est13:AA352603

seq\_documentation\_block: 218 bp mRNA EST 21-APR-1997

LOCUS AA352603

DEFINITION EST760621 Activated T-cells XX Homo sapiens cDNA 5' end similar to similar to major histocompatibility complex, class I, Bw62.3, mRNA sequence.

ACCESSION AA352603

VERSION AA352603.1 GI:2004923

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 218)

AUTHORS Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A., Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C., Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D., Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S., Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr., Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M., Moreno-Palauques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M., Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R., Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y., Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J., Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A., He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K., Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Meissner,P.S., Olsen,H., Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M., Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and Venter,J.C.

TITLE Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence

JOURNAL Nature 377 (6547 Suppl), 3-174 (1995)

MEDLINE 12140200

COMMENT On May 8, 1995 this sequence version replaced gi:800964.

Other\_ESTs: THCI72938

Contact: Kerlavage, AR

Bioinformatics

The Institute for Genomic Research

9712 Medical Center Drive, Rockville, MD 20850 USA

Tel: 3018699056

Fax: 3018699423

Email: arkerlavet@tigr.org

For clone availability, additional sequence and expression

information related to this EST, please check the TIGR Human Gene

Index (<http://www.tigr.org/tdb/hgi/hgi.html>)

Seq primer: M13 Reverse.

FEATURES

Location/Qualifiers

1..218  
 /organism="Homo sapiens"  
 /db\_xref="ATCC (inhost):152802"  
 /db\_xref="taxon:9606"  
 /clone\_lib="Activated T-cells XX"  
 /cell\_type="T-lymphocyte"  
 /dev\_stage="adult"  
 /note="Vector: pBluescript SK-; Site\_1: EcoRI; Site\_2: XhoI"

BASE COUNT 52 a 70 c 62 g 28 t 6 others

ORIGIN



alignment\_scores:  
 Quality: 51.00 Length: 10  
 Ratio: 5.100 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-08-653-294-27 x AA352603 ..  
 Align seg 1/1 to: AA352603 from: 1 to: 218

11 ArgGlusSerLeuArgAsnLeuArgGlyTyr 20  
 |||||  
 39 CGAGAGAGCCTGCGGAACCTGCGGCGTAC 68

seq\_name: gb\_est12:AA294911

seq\_documentation\_block: 259 bp mRNA EST 18-APR-1997  
 LOCUS AA294911  
 DEFINITION EST100074 Pancreas tumor I Homo sapiens cDNA 5' end similar to similar to major histocompatibility complex, class I, Bw62.3, mRNA sequence.

ACCESSION AA294911

VERSION AA294911.1 GI:1947266

## KEYWORDS

EST.

## SOURCE

human.

## ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Hominiidae; Homo.

## REFERENCE

## AUTHORS

Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A., Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C., Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D., Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S., Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr., Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M., Moreno-Palauques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M., Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R., Small,K.V., Springs,T.A., Utterback,T.R., Weidman,J.F., Li,Y., Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J., Dinke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A., He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K., Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Weissner,P.S., Olsen,H., Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M., Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and Venter,J.C.

Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence

Nature 377 (6547 Suppl), 3-174 (1995)

12140200

On Nov 29, 1993 this sequence version replaced gi:430148.

## JOURNAL

## MEDLINE

## COMMENT

Other ESTs: THCI72938

Contact: Kerlavage, AR

Bioinformatics

The Institute for Genomic Research

9712 Medical Center Drive, Rockville, MD 20850 USA

Tel: 3018699056

Fax: 3018699423

Email: arkerlavetigr.org

For clone availability, additional sequence and expression information related to this EST, please check the TIGR Human Gene Index (<http://www.tigr.org/tdb/hgi/hgi.html>)

Seq primer: M13 Reverse.

Location/Qualifiers

1..259

/organism="Homo sapiens"

/db\_xref="ATCC (inhost):190413"

/db\_xref="taxon:9606"

/clone\_lib="Pancreas tumor I"

/dev\_stage="adult"

/note="Organ: pancreas; Vector: pbluescript SK-; Site\_1: EcoRI; Site\_2: XhoI"

BASE COUNT 60 a 87 c 75 g 30 t 7 others

## ORIGIN

alignment\_scores:  
 Quality: 51.00 Length: 10  
 Ratio: 5.100 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-08-653-294-27 x AA294911 ..

Align seg 1/1 to: AA294911 from: 1 to: 259

11 ArgGlusSerLeuArgAsnLeuArgGlyTyr 20

|||||

42 CGAGAGAGCCTGCGGAACCTGCGGCGTAC 71

seq\_name: gb\_est13:AA352960

seq\_documentation\_block: 270 bp mRNA EST 21-APR-1997

LOCUS AA352960

DEFINITION EST61101 Activated T-cells XX Homo sapiens cDNA 5' end similar to similar to major histocompatibility complex, class I, B61, mRNA sequence.

ACCESSION AA352960.1 GI:2005353

VERSION AA352960

KEYWORDS EST.

SOURCE human.

## ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Hominiidae; Homo.

## REFERENCE

## AUTHORS

Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A., Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C., Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D., Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S., Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr., Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M., Moreno-Palauques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M., Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R., Small,K.V., Springs,T.A., Utterback,T.R., Weidman,J.F., Li,Y., Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J., Dinke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A., He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K., Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Weissner,P.S., Olsen,H., Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M., Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and Venter,J.C.

Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence

Nature 377 (6547 Suppl), 3-174 (1995)

12140200

On May 8, 1995 this sequence version replaced gi:801269.

Other ESTs: THCI72938

Contact: Kerlavage, AR

Bioinformatics

The Institute for Genomic Research

9712 Medical Center Drive, Rockville, MD 20850 USA

Tel: 3018699056

Fax: 3018699423

Email: arkerlavetigr.org

For clone availability, additional sequence and expression information related to this EST, please check the TIGR Human Gene Index (<http://www.tigr.org/tdb/hgi/hgi.html>)

Seq primer: M13 Reverse.

Location/Qualifiers

1..270

/organism="Homo sapiens"

/db\_xref="ATCC (inhost):153240"

/db\_xref="taxon:9606"

/clone\_lib="Activated T-cells XX"

/cell\_type="T-lymphocyte"

```

/dev_stage="adult"
/note="vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI"

```

```

BASE COUNT      62 a   80 c   88 g   35 t   5 others
ORIGIN

```

```

alignment_scores:
  Quality:      51.00      Length:      10
  Ratio:        5.100      Gaps:        0
Percent Similarity: 100.000 Percent Identity: 100.000

```

```

alignment_block:
US-08-653-294-27 x AA352960 ..

```

```

Align seg 1/1 to: AA352960 from: 1 to: 270

```

```

11 ArgGluSerLeuArgAsnLeuArgGlyTyr 20
|||||
182 CGAGAGAGCCTNCGGAACCTCGCGGCTAC 211

```

```

seq_name: gb_est8:AA058454

```

```

seq_documentation_block:
LOCUS      AA058454      373 bp      mRNA      EST      01-DEC-1996
DEFINITION z167807.r1 Stratagene colon (#937204) Homo sapiens cDNA clone
IMAGE:509677 5' similar to gb:M24039_cds1 HLA CLASS I
HISTOCOMPATIBILITY ANTIGEN, B-18 B*1801 ALPHA (HUMAN);, mRNA
sequence.

```

```

ACCESSION      AA058454
VERSION        AA058454.1 GI:1551280
KEYWORDS       EST.
SOURCE         human.
ORGANISM       Homo sapiens

```

```

REFERENCE      1 (bases 1 to 373)
AUTHORS        Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
                Chisoe,S., Dietrich,N., Dubuque,T., Favello,A., Gish,W.,
                Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N.,
                Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L.,
                Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Mieg,J.,
                Trevasakis,E., Underwood,K., Wohlmann,P., Waterston,R., Wilson,R.,
                and Marra,M.

```

```

TITLE          Generation and analysis of 280,000 human expressed sequence tags
JOURNAL        Genome Res. 6 (9), 807-828 (1996)
MEDLINE        9704478
COMMENT        On Sep 12, 1996 this sequence version replaced gi:1393357.
                Washington University School of Medicine
                4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
                Tel: 314 286 1800
                Fax: 314 286 1810
                Email: est@watson.wustl.edu
                This clone is available royalty-free through LLNL; contact the
                IMAGE Consortium (info@image.llnl.gov) for further information.
                Insert Length: 1592 Std Error: 0.00
                Seq primer: -28M13 rev2 from Amersham
                High quality sequence stop: 268.
                Location/Qualifiers
                1..373
                /organism="Homo sapiens"
                /db_xref="GDB:3813009"
                /db_xref="taxon:9606"
                /clone="IMAGE:509677"
                /clone_lib="Stratagene colon (#937204)"
                /lab_host="SOLR cells (kanamycin resistant)"
                /note="Organ: colon; Vector: pBluescript SK-; Site_1:
                EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:
                Oligo dT, T-84 colonic epithelial cell line. Average
                insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor
                sequence: 5' GAATTCGGCAGAG 3' -3' adaptor sequence: 5'
                CTCGAGTTTTTTTTTTTTTT 3'"

```

```

FEATURES
  source

```

```

BASE COUNT      73 a   119 c   116 g   57 t   8 others
ORIGIN

```

```

alignment_scores:
  Quality:      51.00      Length:      10
  Ratio:        5.100      Gaps:        0
Percent Similarity: 100.000 Percent Identity: 100.000

```

```

alignment_block:
US-08-653-294-27 x AA058454 ..

```

```

Align seg 1/1 to: AA058454 from: 1 to: 373

```

```

11 ArgGluSerLeuArgAsnLeuArgGlyTyr 20
|||||
229 CGAGAGAGCCTCGGAACCTCGCGGCTAC 258

```

```

seq_name: gb_est11:AA224068

```

```

seq_documentation_block:
LOCUS      AA224068      397 bp      mRNA      EST      19-FEB-1997
DEFINITION zrl1f02.r1 Stratagene hnt neuron (#937233) Homo sapiens cDNA clone
IMAGE:648507 5' similar to gb:L22649_cds1 HLA CLASS I
HISTOCOMPATIBILITY ANTIGEN, B-39 B*3901 ALPHA (HUMAN);, mRNA
sequence.

```

```

ACCESSION      AA224068
VERSION        AA224068.1 GI:1844610
KEYWORDS       EST.
SOURCE         human.
ORGANISM       Homo sapiens

```

```

REFERENCE      1 (bases 1 to 397)
AUTHORS        Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
                Chisoe,S., Dietrich,N., Dubuque,T., Favello,A., Gish,W.,
                Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N.,
                Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L.,
                Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Mieg,J.,
                Trevasakis,E., Underwood,K., Wohlmann,P., Waterston,R., Wilson,R.,
                and Marra,M.

```

```

TITLE          Generation and analysis of 280,000 human expressed sequence tags
JOURNAL        Genome Res. 6 (9), 807-828 (1996)
MEDLINE        9704478
COMMENT        On Nov 29, 1993 this sequence version replaced gi:430429.
                Washington University School of Medicine
                4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
                Tel: 314 286 1800
                Fax: 314 286 1810
                Email: est@watson.wustl.edu
                This clone is available royalty-free through LLNL; contact the
                IMAGE Consortium (info@image.llnl.gov) for further information.
                Seq primer: -28M13 rev1 ET from Amersham
                High quality sequence stop: 295.
                Location/Qualifiers
                1..397
                /organism="Homo sapiens"
                /db_xref="GDB:5586245"
                /db_xref="taxon:9606"
                /clone="IMAGE:648507"
                /clone_lib="Stratagene hnt neuron (#937233)"
                /dev_stage="hnt neurons"
                /lab_host="SOLR (kanamycin resistant)"
                /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
                XhoI; Cloned unidirectionally. Primer: Oligo dT.
                Differentiated, post mitotic hnt neurons. Average insert
                size: 1.5 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5'
                GAATTCGGCAGAG 3' -3' adaptor sequence: 5'
                CTCGAGTTTTTTTTTTTTTT 3'"

```

```

BASE COUNT      77 a   129 c   131 g   56 t   4 others
ORIGIN

```

alignment\_scores:  
 Quality: 51.00 Length: 10  
 Ratio: 5.100 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-08-653-294-27 x AA224068 ..

Align seg 1/1 to: AA224068 from: 1 to: 397

11 ArgGluSerLeuArgAsnLeuArgGlyTyr 20  
 |||||||||||||||||||||||||||||||

297 CGAGAGAGCGCTGCGNAACCTGCGGCTAC 326

seq\_name: gb\_est10:AA160317

seq\_documentation\_block: 405 bp mRNA EST 09-MAR-1998  
 LOCUS AA160317  
 DEFINITION z056c07.r1 Stratagene pancreas (#937208) Homo sapiens cDNA clone  
 IMAGE:590892 5' similar to gb:U11265 HLA CLASS I HISTOCOMPATIBILITY  
 ANTIGEN, B-35 B\*3501 ALPHA (HUMAN);, mRNA sequence.

ACCESSION AA160317

VERSION AA160317.1 GI:1734956

## KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 405)

Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,

Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M.,

Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F.,

Theising, B., White, V., Wylie, T., Waterston, R. and Wilson, R.

WashU-NCI human EST Project

Unpublished (1997)

On Sep 12, 1996 this sequence version replaced gi:1406818.

Contact: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Insert Length: 1671 Std Error: 0.00

Seq primer: -28M13 rev2 from Amersham

High quality sequence stop: 332.

Location/Qualifiers

1..405

/organism="Homo sapiens"

/db\_xref="GDB:4622278"

/db\_xref="taxon:9606"

/clone="IMAGE:590892"

/clone\_lib="Stratagene pancreas (#937208)"

/lab\_host="SOLR cells (kanamycin resistant)"

/note="Organ: pancreas; Vector: pBluescript SK-; Site:1:

ECORI; Site:2: XhoI; Cloned unidirectionally. Primer:

Oligo dT. Pancreatic adenocarcinoma cell line. Average

insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor

sequence: 5' GAATTCGACGAG 3' -3' adaptor sequence: 5'

CTCAGATTTTTTTTTTTT 3'

BASE COUNT 80 a 132 c 132 g 58 t 3 others

ORIGIN

## alignment\_scores:

Quality: 51.00 Length: 10

Ratio: 5.100 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-08-653-294-27 x AA160317 ..

Align seg 1/1 to: AA160317 from: 1 to: 405

11 ArgGluSerLeuArgAsnLeuArgGlyTyr 20

|||||||||||||||||||||||||||||

281 CGAGAGAGCGCTGCGNAACCTGCGGCTAC 310

seq\_name: gb\_est9:AA100680

seq\_documentation\_block: 419 bp mRNA EST 31-JUL-1997

LOCUS AA100680

DEFINITION z190b02.r1 Stratagene colon (#937204) Homo sapiens cDNA clone

IMAGE:511851 5' similar to gb:M24039\_cds1 HLA CLASS I

HISTOCOMPATIBILITY ANTIGEN, B-18 B\*1801 ALPHA (HUMAN);, mRNA

sequence.

ACCESSION AA100680

VERSION AA100680.1 GI:1646981

KEYWORDS

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 419)

Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B.,

Chisoe, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W.,

Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N.,

Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L.,

Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J.,

Trevaskis, E., Underwood, K., Wohldmann, P., Waterston, R., Wilson, R.

and Marra, M.

Generation and analysis of 280,000 human expressed sequence tags

Genome Res. 6 (9), 807-828 (1996)

97044478

Contact: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Insert Length: 1534 Std Error: 0.00

Seq primer: -28M13 rev2 from Amersham

High quality sequence stop: 275.

Location/Qualifiers

1..419

/organism="Homo sapiens"

/db\_xref="GDB:3844346"

/db\_xref="taxon:9606"

/clone="IMAGE:511851"

/clone\_lib="Stratagene colon (#937204)"

/lab\_host="SOLR cells (kanamycin resistant)"

/note="Organ: colon; Vector: pBluescript SK-; Site:1:

ECORI; Site:2: XhoI; Cloned unidirectionally. Primer:

Oligo dr. T-84 colonic epithelial cell line. Average

insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor

sequence: 5' GAATTCGACGAG 3' -3' adaptor sequence: 5'

CTCAGATTTTTTTTTTTT 3'

BASE COUNT 85 a 133 c 133 g 63 t 5 others

ORIGIN

## alignment\_scores:

Quality: 51.00 Length: 16

Ratio: 3.643 Gaps: 1

Percent Similarity: 87.500 Percent Identity: 75.000

## alignment\_block:

US-08-653-294-27 x AA100680 ..

Align seg 1/1 to: AA100680 from: 1 to: 419

5 AsnArgLeuSerGluArgArgGluSerLeuArgAsnLeuArgGlyTyr 20

```

:::|||||:|||||
225 CACAGACTACGAGAGA.....AGCCTGGGAACCTGCGGGCTAC 266
seq_name: gb_est7:W40489

seq_documentation_block:
LOCUS      W40489      427 bp      mRNA      EST      20-MAY-1996
DEFINITION zc84b01.r1 Pancreatic Islet Homo sapiens cDNA clone IMAGE:328969 5'
            similar to gb:U11265 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-35
            B*3501 ALPHA (HUMAN);, mRNA sequence.
ACCESSION  W40489
VERSION     W40489.1  GI:1324496
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 427)
AUTHORS     Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B.,
            Chisoe, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W.,
            Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N.,
            Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L.,
            Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J.,
            Trevas, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R.,
            and Marra, M.
            Generation and analysis of 280,000 human expressed sequence tags
            Genome Res. 6 (9), 807-828 (1996)
JOURNAL     97044478
MEDLINE
COMMENT     On Jan 25, 1995 this sequence version replaced gi:637865.
            Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LNL; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: mob.REGA+T
            High quality sequence stop: 397.
FEATURES
            Location/Qualifiers
            1..427
            /organism="Homo sapiens"
            /db_xref="GDB:1263173"
            /db_xref="taxon:9606"
            /clone="IMAGE:328969"
            /clone_lib="Pancreatic Islet"
            /tissue_type="pancreatic islet"
            /lab_host="SOLR cells (kanamycin resistant)"
            /note="Organ: pancreas; Vector: pBluescript SK-; Site_1:
            EcoRI; Site_2: XhoI; Reference: Hum Mol Gen 2, 1795 (1993)
            Takeda et al. Cloned unidirectionally. Primer: Oligo dt.
            -5' adaptor sequence: 5' GAATTCGACGAG 3' -3' adaptor
            sequence: 5' CTCGAGTTTCTTTTCTTTT 3"
            3 others
BASE COUNT      87 a      141 c      134 g      62 t
ORIGIN
alignment_scores:
            Quality: 51.00      Length: 10
            Ratio: 5.100      Gaps: 0
            Percent Similarity: 100.000      Percent Identity: 100.000
alignment_block:
US-08-653-294-27 x W40489
Align seg 1/1 to: W40489 from: 1 to: 427
11 ArgGluSerLeuArgAsnLeuArgGlyTyr 20
|||||
279 CGAGAGACCTCGGAACCTCGCGGCTAC 308
seq_name: gb_est12:AA310808

```

```

seq_documentation_block:
LOCUS      AA310808      443 bp      mRNA      EST      19-APR-1997
DEFINITION EST181593 Jurkat T-cells V Homo sapiens cDNA 5' end similar to
            similar to major histocompatibility complex, class I, B0704, mRNA
            sequence.
ACCESSION  AA310808
VERSION     AA310808.1  GI:1963136
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 443)
AUTHORS     Adams, M.D., Kerlavage, A.R., Fleischmann, R.D., Fuldner, R.A.,
            Bult, C.J., Lee, N.H., Kirkness, E.F., Weinstock, K.G., Gocayne, J.D.,
            White, O., Sutton, G., Blake, J.A., Brandon, R.C., Man, Wai, C.,
            Clayton, R.A., Cline, T.R., Cotton, M.D., Earle-Hughes, J., Fine, L.D.,
            Fitzgerald, L.M., Fitzhugh, W.M., Fritchman, J.L., Geoghagen, N.S.,
            Glodak, A., Gnehm, C.L., Hanna, M.C., Hedblom, E., Hinkle, P.S., Jr.,
            Kelley, J.M., Kelley, J.C., Liu, L.-I., Marmaros, S.M., Merrick, J.M.,
            Morano, Palanques, R.F., McDonald, L.A., Nguyen, D.T., Pelligrino, S.M.,
            Phillips, C.A., Ryder, S.E., Scott, J.L., Saudek, D.M., Shirley, R.,
            Small, K.V., Spriggs, T.A., Utterback, T.R., Weidman, J.F., Li, Y.,
            Bednarek, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.J.,
            Dimke, D., Feng, D.-F., Ferrie, A., Fischer, C., Hastings, G.A.,
            He, W.W., Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K.,
            Kozak, D.L., Kunsch, C., Hungjun, J., Li, H., Meissner, P.S., Olsen, H.,
            Raymond, L., Wei, Y.F., Wing, J., Xu, C., Yu, G.L., Ruben, S.M.,
            Dillion, P.J., Fannon, M.R., Rosen, C.A., Haseltine, W.A., Fields, C.,
            Fraser, C.M. and Venter, J.C.
            Initial assessment of human gene diversity and expression patterns
            based upon 83 million nucleotides of cDNA sequence
            Nature 377 (6547 Suppl), 3-174 (1995)
JOURNAL     12140200
MEDLINE
COMMENT     On Sep 12, 1996 this sequence version replaced gi:1397854.
            Other ESTs: TNC180721
            Contact: Kerlavage, AR
            Bioinformatics
            The Institute for Genomic Research
            9712 Medical Center Drive, Rockville, MD 20850 USA
            Tel: 3018699056
            Fax: 3018699423
            Email: arkerlavet@tigr.org
            For clone availability, additional sequence and expression
            information related to this EST, please check the TIGR Human Gene
            Index (http://www.tigr.org/tldb/hgi/hgi.html)
            Seq primer: M13 Reverse.
FEATURES
            Location/Qualifiers
            1..443
            /organism="Homo sapiens"
            /db_xref="ATCC (inhost):156811"
            /db_xref="taxon:9606"
            /clone_lib="Jurkat T-cells V"
            /cell_type="T-lymphocyte"
            /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
            XhoI"
BASE COUNT      86 a      144 c      143 g      64 t      6 others
ORIGIN
alignment_scores:
            Quality: 51.00      Length: 10
            Ratio: 5.100      Gaps: 0
            Percent Similarity: 100.000      Percent Identity: 100.000
alignment_block:
US-08-653-294-27 x AA310808
Align seg 1/1 to: AA310808 from: 1 to: 443
11 ArgGluSerLeuArgAsnLeuArgGlyTyr 20
|||||
292 CGAGAGACCTTCGGAACCTCGCGGCTAC 321

```

```

seq_name: gb_est9:C18310
seq_documentation_block: 475 bp mRNA EST 02-OCT-1996
LOCUS C18310 Human placenta cDNA (TFujiwara) Homo sapiens cDNA clone
DEFINITION GEN:560D07 5', mRNA sequence.
ACCESSION C18310
VERSION C18310.1 GI:1579912
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 475)
AUTHORS Fujiwara, T., Hirano, H., Hishigaki, H., Horie, M., Kawai, A., Kuga, Y.,
Kyushiki, H., Nagata, S., Okuno, S., Ozaki, K., Shimizu, F., Shimada, Y.,
Shinomiya, H., Suzuki, M., Takaichi, A., Takeda, S., Watanabe, T.,
Maekawa, H., Nakamura, Y. and Takahashi, E.
TITLE Otsuka cDNA project
JOURNAL Unpublished (1996)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1393837.
Contact: Tsutomu Fujiwara
Otsuka GEN Research Institute
Otsuka Pharmaceutical Co., Ltd
463-10 Kagasuno Kawauchi-cho, Tokushima, Tokushima, 771-01 Japan
Tel: 0886-65-2888
Fax: 0886-37-1035.
FEATURES
Location/Qualifiers
source
1..475
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="GEN-560D07"
/clone_lib="Human placenta cDNA (TFujiwara)"
/tissue_type="Placenta" 68 t 3 others
BASE COUNT 89 a 161 c 154 g 68 t
ORIGIN

alignment_scores:
Quality: 51.00 Length: 10
Ratio: 5.100 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-27 x C18310 ..
Align seg 1/1 to: C18310 from: 1 to: 475
11 ArgGluSerLeuArgAsnLeuArgGlyTyr 20
|||||
299 CGAGAGAGCCTGCGAACCTGCGCGCTAC 328

seq_name: gb_est17:AA663896
seq_documentation_block: 479 bp mRNA EST 15-DEC-1997
LOCUS AA663896
DEFINITION ae74d01.s1 Stratagene schizo brain S11 Homo sapiens cDNA clone
IMAGE:969889 3' similar to gb:M28203 HLA CLASS I HISTOCOMPATIBILITY
ANTIGEN, B-62 B*1504 ALPHA (HUMAN);, mRNA sequence.
ACCESSION AA663896
VERSION AA663896.1 GI:2617887
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 479)
AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,
Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M.,
Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F.,
Theising, B., White, Y., Wyllie, T., Waterston, R. and Wilson, R.
TITLE WashU-NCI human EST project

```

```

JOURNAL Unpublished (1997)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1394858.
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL: contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Possible reversed clone: polyt not found
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 430.
FEATURES
Location/Qualifiers
source
1..479
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:969889"
/clone_lib="Stratagene schizo brain S11"
/sex="male"
/tissue_type="schizophrenic brain S-11 frontal lobe"
/dev_stage="34 years old"
/lab_host="SOLR (kanamycin resistant)"
/notes="vector: Bluescript SK-; Site_1: EcoRI; Library
constructed from S-11 frontal lobe, male, 34 years old,
50% caucasian, 50% Aleutian. Schizophrenic suicide.
Random primed into EcoRI site of ZAP II Vector. Mass
excised. Custom library. Avg insert length 1.4kb.
Material obtained by Johnston N., Torrey, E.F., Yolken R.,
and the Stanley Neuropathology Consortium - Analysis of
RNAs from the Brains of Individuals with Psychiatric
Diseases (Unpublished) Stanley Neurovirology Laboratory,
Johns Hopkins School of Medicine, Baltimore MD."
BASE COUNT 91 a 165 c 158 g 64 t 1 Others
ORIGIN

alignment_scores:
Quality: 51.00 Length: 10
Ratio: 5.100 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-27 x AA663896 ..
Align seg 1/1 to: AA663896 from: 1 to: 479
11 ArgGluSerLeuArgAsnLeuArgGlyTyr 20
|||||
283 CGAGAGAGCCTGCGAACCTGCGCGCTAC 312

seq_name: gb_est35:AL039796
seq_documentation_block: 503 bp mRNA EST 29-SEP-1999
LOCUS AL039796
DEFINITION DKZPD434B1912 r1 434 (synonym: htes3) Homo sapiens cDNA clone
IMAGE:969889 3' similar to gb:M28203 HLA CLASS I HISTOCOMPATIBILITY
ANTIGEN, B-62 B*1504 ALPHA (HUMAN);, mRNA sequence.
ACCESSION AL039796
VERSION AL039796.1 GI:5408804
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 503)
AUTHORS Duesterhoeft, A., Lauber, J., Mewes, H.W., Gassenhuber, J. and
Wiemann, S.
TITLE EST (Duesterhoeft, et al.)
JOURNAL Unpublished (1999)
COMMENT On Jun 2, 1999 this sequence version replaced gi:4967270.
Contact: Duesterhoeft A
MIPS
Am Klopferspitz 18a D-82152 Martinsried, Germany

```

This is the 5' sequence of the clone insert  
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
Sponsored by the German Cancer Research Consortium of the  
German Genome Project.

## FEATURES

**source**

```

BASE COUNT
ORIGIN
96 a 172 c 165 g 68 t 2 others
/note="Vector: pspori1; Site_1: NotI; Site_2: SalI"

```

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2118100"
/clone_lib="NCI_CGAP CLL1"
/tissue_type="B-cell, chronic lymphocytic le
/lab_host="DH10B"
/notes="vector: p77r3d-Pac (Pharmacia) with
polylinker; Site_1: Not I; Site_2: Eco RI;
was primed with a Not I - oligo(dT) primer;
TGTTACCAATCTGAAGTGGAGCGCGCATTCGCTTTT
T 3']; double-stranded cDNA was ligated to
adaptors (Pharmacia), digested with Not I
the Not I and Eco RI sites of the modified
library is normalized, and was constructed
Soares and M.Fatima Bonaldo."
BASE COUNT      122 a 138 c 202 g 120 t      1 others
ORIGIN

alignment_scores:
    Quality:      51.00
    Ratio:        3.400
    Percent Similarity: 83.333    Percent Identity: 61.111

alignment_block:
US-08-653-294-27 x AI524732 ..

Align seg 1/1 to: AI524732 from: 1 to: 583

2 GlyArgLeuAsnArgLeuSerGluArgArgGluSerLeuArgAspLeuAr 18
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
522 GGCGCGTTGAGCGCGNTTTGCAGTCGGAGGGCGAGCCTTAAGAGGAGCCCA 571

18 gGly 19
:||||
572 GGGA 575

```

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2000, 04:05:43 ; Search time 133.56 Seconds  
(without alignments)  
3.547 Million cell updates/sec

Title: US-08-653-294-28

Perfect score: 99

Sequence: 1 YRLATLNERENLRALRY 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 188963 seqs, 23686106 residues

Total number of hits satisfying chosen parameters; 188963

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : A\_Geneseq\_36.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	99	100.0	20	1 R92908	HLA-B*2702 CTL modu
2	99	100.0	20	1 W33791	Peptide B2702.84-7
3	94	94.9	20	1 R92910	HLA-B*2702 CTL modu
4	94	94.9	20	1 W33793	Peptide B2702.84-7
5	93	93.9	20	1 R92907	HLA-B*2702 CTL modu
6	93	93.9	20	1 R95428	HLA-B*2702 84-75-84
7	93	93.9	20	1 W33778	Immunomodulating d
8	88	88.9	20	1 R92909	HLA-B*2702 CTL modu
9	88	88.9	20	1 W33792	Peptide B2702.84-7
10	80	80.8	20	1 R95430	HLA-B*2702 84-75/7
11	50	50.5	25	1 R48286	Peptide fragment o
12	50	50.5	25	1 R83093	HLAB38 CTL modul
13	50	50.5	25	1 R95422	HLAB38.6084. Comps
14	50	50.5	184	1 Y06801	Peptide Seq ID No:
15	50	50.5	362	1 R03142	Sequence of HLA-Bw
16	50	50.5	362	1 R03144	Sequence of HLA-B5
17	50	50.5	362	1 R12463	HLA-B*53 exon. HLA
18	49.5	50.0	20	1 R92911	HLA-B*2702 CTL modu
19	49.5	50.0	20	1 W33779	Immunomodulating d
20	49	49.5	10	1 R41208	Peptide fragment o
21	49	49.5	10	1 R83062	HLA-B*2702 CTL modu
22	49	49.5	10	1 R95413	Alphal-helix of HL
23	49	49.5	10	1 W07512	T-cell modulating
24	49	49.5	10	1 W47265	Immunomodulatory p
25	49	49.5	10	1 W33784	Peptide B2702.75-8
26	49	49.5	15	1 R92912	HLA-B*2702 CTL modu
27	49	49.5	15	1 W33795	Peptide B2702.70-8
28	49	49.5	25	1 R41205	Peptide fragment o
29	49	49.5	25	1 R83090	HLA-B*2702 CTL modu
30	49	49.5	25	1 R95416	HLA-B*2702 60-84. C
31	49	49.5	25	1 W33794	Peptide B2702.60-8
32	44	44.4	10	1 R83094	HLA-B*2702 CTL modu
33	44	44.4	10	1 R83095	HLA-B*2702 CTL modu
34	44	44.4	10	1 R83096	HLA-B*2702 CTL modu

35 44 44.4 10 1 R95425 HLA-B\*2702.75-84 (D)  
36 44 44.4 10 1 R95426 HLA-B\*2702.75-84 (T)  
37 44 44.4 10 1 W07513 T-cell modulating  
38 44 44.4 10 1 W47266 Immunomodulatory p  
39 44 44.4 10 1 W47267 Immunomodulatory p  
40 44 44.4 10 1 W47269 Immunomodulatory p  
41 44 44.4 10 1 W33788 Peptide B2702.75-8  
42 44 44.4 10 1 W33787 Peptide B2702.75-8  
43 44 44.4 10 1 W33789 Peptide B2702.75-8  
44 42 42.4 20 1 R92913 HLA-B7 CTL modul  
45 42 42.4 20 1 R95415 HLA-B7.84-75-84 Pa

#### ALIGNMENTS

RESULT 1

R92908 R92908 standard; peptide; 20 AA.  
AC R92908;  
DT 16-MAY-1996 (first entry)  
DE HLA-B\*2702 CTL modulating peptide (B2702.84-75(T)/75-84).  
KW Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;  
KW immunosuppressant; graft versus host disorder; transplantation; therapy;  
KW class I MHC; HLA-B\*2702.  
OS Synthetic.  
PN W09526979-A1.  
PD 12-OCT-1995.  
PF 05-APR-1995; U04349.  
PR 05-APR-1994; US-222851.  
PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
PI Clayberger C, Krensky AM, Farham P;  
DR WPI: 95-358582/46.  
PT Extension of acceptance period of transplants from MHC unmatched  
PT donor hosts - using Class I B75-84 MHC antigen of the recipient  
PT host  
FS Example 15; Page 36; 80pp; English.  
CC R83061-R83085, R83090-R83096 and R92907-R92914 represent fragments of  
CC class I major histocompatibility complex (MHC) antigens. This sequence  
CC is an inverted dimer of residues 75-84 of the alpha-1 domain of the class  
CC I MHC HLA-B\*2702. These sequences can be used to extend the period of  
CC acceptance by a recipient of a transplant from an MHC unmatched donor.  
CC The peptides are administered to a patient in conjunction with a  
CC subtherapeutic amount of an immunosuppressant. This is administered to  
CC the patient for a limited period of time (compared to the lifetime  
CC administration for current treatments). The peptides particularly  
CC modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)  
CC of the patient.  
SQ Sequence 20 AA;

Query Match 100.0%; Score 99; DB 1; Length 20;

Best Local Similarity 100.0%; Pred. No. 4.1e-10; Mismatches 0; Indels 0; Gaps 0;  
Matches 20; Conservative 0;

Qy 1 YRLATLNERENLRALRY 20

Db 1 YRLATLNERENLRALRY 20

RESULT 2

W33791 W33791 standard; peptide; 20 AA.

AC W33791;  
DT 19-JUN-1998 (first entry)  
DE Peptide B2702.84-75r/75-84 tested for immunomodulating activity.  
KW Immunomodulating dimer; immunosuppressant drug; CTL activation;  
KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;  
KW rejection.  
OS Synthetic.  
PN W09744351-A1.  
PD 27-NOV-1997.

```

PF 22-MAY-1997; U08689.
PR 24-MAY-1996; US-653294.
PI (STRD ) UNIV LELAND STANFORD JUNIOR.
PI Beulow R, Clayberger C, Krensky AM;
DR WPI: 98-086530/08.
DR New immunomodulating dimer peptide(s) - based on a Class I HLA-B
PT alpha-1 domain, used for preventing rejection of transplants or
PT treating autoimmune diseases
PS Example 1; Page 19; 41pp; English.
CC Peptides W33784-98 and W33778-9 were assayed for their immunomodulating
CC activity. A peptide-type compound or variant is claimed which has
CC immunomodulating activity, including the N-terminal acylated and/or
CC C-terminal amidated or esterified forms of up to 60 amino acids, where
CC the peptide-type compound comprises the formula: A-B, where A, B =
CC (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or V; aa77 =
CC D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a hydrophobic or
CC small amino acid; aa82 = R or L; aa83 = G or R; and aa represents amino
CC acid. The sequence in the brackets may optionally be absent or truncated
CC at any peptide type bond within the brackets. The compounds comprise
CC amino acid sequences related to a Class I HLA-B alpha1 domain (positions
CC 79-84). They can be used to inhibit cytotoxic T-lymphocytes (CTL) from
CC undesirably attacking cells in a host or in vitro. They can also be
CC used in combination with antigenic peptides or proteins of interest to
CC activate CTLs. They can also inhibit the proliferation of T cells in
CC response to anti-CD3. The peptide can be used for preventing rejection
CC of transplants or for treating autoimmune diseases, e.g. diabetes,
CC rheumatoid arthritis and lupus erythematosus. The products can also be
CC used for detection and diagnosis.
CC Sequence 20 AA;
SQ

Query Match 100.0%; Score 99; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 4.1e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLATRLNERENLRALRY 20
DB 1 YRLATRLNERENLRALRY 20
|||||
RESULT 3
R92910
ID R92910 standard; peptide; 20 AA.
AC R92910;
DT 16-MAY-1996 (first entry)
DE HLA-B2702 CTL modulating peptide (B2702.84-75(T)/75-84(T)).
KW Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;
KW immunosuppressant; graft versus host disorder; transplantation; therapy;
KW Class I MHC; HLA-B2702.
OS Synthetic.
PN W09526979-A1.
PD 12-OCT-1995.
PF 05-APR-1995; U04349.
PR 05-APR-1994; US-222851.
PI (STRD ) UNIV LELAND STANFORD JUNIOR.
PI Clayberger C, Krensky AM, Parham P;
DR WPI: 95-358582/46.
DR Extension of acceptance period of transplants from MHC unmatched
PT donor hosts - using Class I B7-84 MHC antigen of the recipient
PT host
PS Example 15; Page 36; 80pp; English.
CC R83061-R83085, R83090-R83096 and R92907-R92913 represent fragments of
CC Class I major histocompatibility complex (MHC) antigens. This sequence
CC is an inverted dimer of residues 75-84 of the alpha-1 domain of the class
CC I MHC HLA-B2702. These sequences can be used to extend the period of
CC acceptance by a recipient of a transplant from an MHC unmatched donor.
CC The peptides are administered to a patient in conjunction with a
CC subtherapeutic amount of an immunosuppressant. This is administered to
CC the patient for a limited period of time (compared to the lifetime
CC administration for current treatments). The peptides particularly
CC modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)
CC of the patient.
CC Sequence 20 AA;
SQ

Query Match 94.9%; Score 94; DB 1; Length 20;
Best Local Similarity 95.0%; Pred. No. 2.7e-09;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 YRLATRLNERENLRALRY 20
DB 1 YRLATRLNERENLRALRY 20
|||||
RESULT 5
R92907
ID R92907 standard; peptide; 20 AA.
AC R92907;
DT 16-MAY-1996 (first entry)
DE HLA-B2702 CTL modulating peptide (B2702.84-75(T)/75-84(T)).
KW Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;
KW immunosuppressant; graft versus host disorder; transplantation; therapy;
KW class I MHC; HLA-B2702.
SQ

Query Match 94.9%; Score 94; DB 1; Length 20;
Best Local Similarity 95.0%; Pred. No. 2.7e-09;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 YRLATRLNERENLRALRY 20
DB 1 YRLATRLNERENLRALRY 20
|||||
RESULT 4
ID W33793 standard; peptide; 20 AA.
AC W33793;
DT 19-JUN-1998 (first entry)
DE Peptide B2702.84-75(T)/75-84(T) tested for immunomodulating activity.
KW immunomodulating dimer; immunosuppressant drug; CTL activation;
KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;
KW rejection.
OS Synthetic.
OS Homo sapiens.
PN W09744351-A1.
PD 27-NOV-1997.
PF 22-MAY-1997; U08689.
PR 24-MAY-1996; US-653294.
PI (STRD ) UNIV LELAND STANFORD JUNIOR.
PI Beulow R, Clayberger C, Krensky AM;
DR WPI: 98-086530/08.
DR New immunomodulating dimer peptide(s) - based on a Class I HLA-B
PT alpha-1 domain, used for preventing rejection of transplants or
PT treating autoimmune diseases
PS Example 1; Page 19; 41pp; English.
CC Peptides W33784-98 and W33778-9 were assayed for their immunomodulating
CC activity. A peptide-type compound or variant is claimed which has
CC immunomodulating activity, including the N-terminal acylated and/or
CC C-terminal amidated or esterified forms of up to 60 amino acids, where
CC the peptide-type compound comprises the formula: A-B, where A, B =
CC (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or V; aa77 =
CC D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a hydrophobic or
CC small amino acid; aa82 = R or L; aa83 = G or R; and aa represents amino
CC acid. The sequence in the brackets may optionally be absent or truncated
CC at any peptide type bond within the brackets. The compounds comprise
CC amino acid sequences related to a Class I HLA-B alpha1 domain (positions
CC 79-84). They can be used to inhibit cytotoxic T-lymphocytes (CTL) from
CC undesirably attacking cells in a host or in vitro. They can also be
CC used in combination with antigenic peptides or proteins of interest to
CC activate CTLs. They can also inhibit the proliferation of T cells in
CC response to anti-CD3. The peptide can be used for preventing rejection
CC of transplants or for treating autoimmune diseases, e.g. diabetes,
CC rheumatoid arthritis and lupus erythematosus. The products can also be
CC used for detection and diagnosis.
CC Sequence 20 AA;
SQ

```



OS Synthetic.  
 PN W09526979-A1.  
 PD 12-OCT-1995.  
 PF 05-APR-1995.  
 PR 05-APR-1995; U04349.  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI Clayberger C, Krensky AM, Parham P;  
 DR WPI; 95-358582/46.  
 PT Extension of acceptance period of transplants from MHC unmatched  
 PT donor hosts - using Class I B75-84 MHC antigen of the recipient  
 PT host  
 PS Example 15; Page 36; 80pp; English.  
 CC R83061-R83085, R83090-R83096 and R92907-R92914 represent fragments of  
 CC class I major histocompatibility complex (MHC) antigens. This sequence  
 CC is an inverted dimer of residues 75-84 of the alpha-1 domain of the class  
 CC I MHC HLA-B\*2702. These sequences can be used to extend the period of  
 CC acceptance by a recipient of a transplant from an MHC unmatched donor.  
 CC The peptides are administered to a patient in conjunction with a  
 CC subtherapeutic amount of an immunosuppressant. This is administered to  
 CC the patient for a limited period of time (compared to the lifetime  
 CC administration for current treatments). The peptides particularly  
 CC modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)  
 CC of the patient.  
 CC Sequence 20 AA;

Query Match 93.9%; Score 93; DB 1; Length 20;  
 Best Local Similarity 95.0%; Pred. No. 3.9e-09;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 YRLATRLNERENRLALRY 20  
 DB 1 YRLAIRLNERENRLALRY 20

RESULT 6  
 ID R95428 standard; peptide: 20 AA.  
 AC R95428;  
 DT 12-NOV-1996 (first entry)  
 DE HLA-B\*2702 84-75-84 palindromic.  
 KW HLA; p74; alpha1-helix; human-leucocyte-associated antigen; inhibitor;  
 KW T-cell lysate; membrane protein; mammal; heat shock protein; Hsc70; APC;  
 KW B cell; calcium influx; cytotoxic T lymphocyte; CTL; differentiation;  
 KW cytolysis; antigen presenting cell.  
 OS Synthetic.  
 PN W09513288-A1.  
 PD 18-MAY-1995.  
 PF 10-NOV-1994; U12985.  
 PR 10-NOV-1993; US-150493.  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI Clayberger C, Krensky AM;  
 DR WPI; 95-194027/25.  
 PT Compsns. comprising lymphoid surface membrane proteins - which may  
 PT inhibit cytolytic activity and differentiation of CTLs.  
 PS Example; Page 12; 29pp; English.  
 CC R95413, and R95415-R95431 represent palindromes and fragments of  
 CC human-leucocyte-associated antigens. This sequence represents the  
 CC HLA-B\*2702 84-75-84 palindromic. These sequences can be used to isolate  
 CC the protein p74 from a T-cell lysate. p74 is a T-cell surface membrane  
 CC protein associated with T-cell activation in mammalian T-cells, and is  
 CC also immunologically cross reactive with the heat shock protein Hsc70.  
 CC p74 is found in a limited number of cell types, but is particularly  
 CC expressed on B and T cells. p74 can be isolated by lysis of a suitable  
 CC cell with an amphoteric detergent, and then passed through an affinity  
 CC column containing a covalently bound HLA-B\*2702 palindromic peptide.  
 CC Compositions comprising the extracellular fragment of p74 combined with  
 CC HLA-B\*2702 60-84 (see R95416), induces calcium influx, and inhibits  
 CC cytotoxic T lymphocyte (CTL) differentiation or cytolysis. Candidate  
 CC compounds can be screened for their effect on the cytolytic activity of  
 CC T-cells, by combining them with the extracellular portion of p74 and  
 CC determining the amount of binding between the candidate compound and p74.  
 CC Modulation of CTL activity can be inhibited in a cellular composition

CC containing T-cells and antigen presenting cells (APCs), by adding to the  
 CC mix the extracellular portion of p74, in an amount sufficient to compete  
 CC with p74 for the binding of the p74 ligand.  
 SQ Sequence 20 AA;

Query Match 93.9%; Score 93; DB 1; Length 20;  
 Best Local Similarity 95.0%; Pred. No. 3.9e-09;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 YRLATRLNERENRLALRY 20  
 DB 1 YRLAIRLNERENRLALRY 20

RESULT 7  
 W33778  
 ID W33778 standard; peptide: 20 AA.  
 AC W33778;  
 DT 19-JUN-1998 (first entry)  
 DE Immunomodulating dimer peptide #1.  
 KW Immunomodulating dimer; immunosuppressant drug; CTL activation;  
 KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;  
 KW rejection.  
 OS Synthetic.  
 PN Homo sapiens.  
 PD W09744351-A1.  
 PF 27-NOV-1997.  
 PR 22-MAY-1997; U08689.  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI Beulow R, Clayberger C, Krensky AM;  
 DR WPI; 98-086530/08.  
 PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B  
 PT alpha-1 domain, used for preventing rejection of transplants or  
 PT treating autoimmune diseases  
 PS Claim 16; Page 35; 41pp; English.  
 CC This sequence represents a specifically claimed immunomodulating  
 CC dimer peptide of the invention. A peptide-type compound or variant is  
 CC claimed which has immunomodulating activity, including the N-terminal  
 CC acylated and/or C-terminal amidated or esterified forms of up to 60  
 CC amino acids, where the peptide-type compound comprises the formula: A-B,  
 CC where A, B = (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or  
 CC V; aa77 = D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a  
 CC hydrophobic or small amino acid; aa82 = R or L; aa83 = G or R; and aa  
 CC represents amino acid. The sequence in the brackets may optionally be  
 CC absent or truncated at any peptide type bond within the brackets. The  
 CC compounds comprise amino acid sequences related to a Class I HLA-B  
 CC alfa1 domain (positions 79-84). They can be used to inhibit cytotoxic  
 CC T-lymphocytes (CTL) from undesirably attacking cells in a host or in  
 CC vitro. They can also be used in combination with antigenic peptides or  
 CC proteins of interest to activate CTLs. They can also inhibit the  
 CC proliferation of T cells in response to anti-CD3. The peptide can be  
 CC used for preventing rejection of transplants or for treating autoimmune  
 CC diseases, e.g. diabetes, rheumatoid arthritis and lupus erythematosus.  
 CC The products can also be used for detection and diagnosis.  
 CC Sequence 20 AA;

Query Match 93.9%; Score 93; DB 1; Length 20;  
 Best Local Similarity 95.0%; Pred. No. 3.9e-09;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 YRLATRLNERENRLALRY 20  
 DB 1 YRLAIRLNERENRLALRY 20

RESULT 8  
 R92909  
 ID R92909 standard; peptide: 20 AA.  
 AC R92909;  
 DT 16-MAY-1996 (first entry)

DE HLA-B2702 CTL modulating peptide (B2702.84-75/75-84(T)).  
 KW Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;  
 KW immunosuppressant; graft versus host disorder; transplantation; therapy;  
 KW class I MHC; HLA-B2702.  
 OS Synthetic.  
 PN W09526979-A1.  
 PD 12-OCT-1995.  
 PF 05-APR-1995; U04349.  
 PR 05-APR-1994; US-222851.  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI Clayberger C, Krensky AM, Parham P;  
 DR WPI: 95-358562/46.  
 PT Extension of acceptance period of transplants from MHC unmatched  
 PT donor hosts - using Class I B75-84 MHC antigen of the recipient  
 PT host  
 PS Example 15; Page 36; 80pp; English.  
 CC R3061-R83085, R83090-R83096 and R92907-R92913 represent fragments of  
 CC class I major histocompatibility complex (MHC) antigens. This sequence  
 CC is an inverted dimer of residues 75-84 of the alpha-1 domain of the class  
 CC I MHC HLA-B2702. These sequences can be used to extend the period of  
 CC acceptance by a recipient of a transplant from an MHC unmatched donor.  
 CC The peptides are administered to a patient in conjunction with a  
 CC subtherapeutic amount of an immunosuppressant. This is administered to  
 CC the patient for a limited period of time (compared to the lifetime  
 CC administration for current treatments). The peptides particularly  
 CC modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)  
 CC of the patient.  
 SQ Sequence 20 AA;

Query Match 88.9%; Score 88; DB 1; Length 20;  
 Best Local Similarity 90.0%; Pred. No. 2.6e-08;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 YRLATRLNERENLRALRY 20  
 ||||| ||||| ||||| |||||  
 Db 1 YRLAIRLNERENLRALRY 20

## RESULT 9

ID W33792 standard; peptide; 20 AA.  
 AC W33792:  
 DT 19-JUN-1998 (first entry)  
 DE Peptide B2702.84-75/75-84T tested for immunomodulating activity.  
 KW Immunomodulating dimer; immunosuppressant drug; CTL activation;  
 KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;  
 KW rejection.  
 OS Synthetic.  
 OS Homo sapiens.  
 PN W09744351-A1.  
 PD 27-NOV-1997; U08689.  
 PF 22-MAY-1997; U08689.  
 PR 24-MAY-1996; US-653294.  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI Beulow R, Clayberger C, Krensky AM;  
 DR WPI: 98-086530/08.  
 PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B  
 PT alpha-1 domain, used for preventing rejection of transplants or  
 PT treating autoimmune diseases  
 PS Example 1; Page 19; 41pp; English.  
 CC Peptides W33784-98 and W33778-9 were assayed for their immunomodulating  
 CC activity. A peptide-type compound or variant is claimed which has  
 CC C-terminal amidated or esterified forms of up to 60 amino acids, where  
 CC the peptide-type compound comprises the formula: A-B, where A, B =  
 CC (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or V; aa77 =  
 CC D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a hydrophobic or  
 CC small amino acid; aa82 = R or L; aa83 = G or R; and aa represents amino  
 CC acid. The sequence in the brackets may optionally be absent or truncated  
 CC at any peptide type bond within the brackets. The compounds comprise  
 CC amino acid sequences related to a Class I HLA-B alpha domain (positions  
 CC 79-84). They can be used to inhibit cytotoxic T-lymphocytes (CTL) from

CC undesirably attacking cells in a host or in vitro. They can also be  
 CC used in combination with antigenic peptides or proteins of interest to  
 CC activate CTLs. They can also inhibit the proliferation of T cells in  
 CC response to anti-CD3. The peptide can be used for preventing rejection  
 CC of transplants or for treating autoimmune diseases, e.g. diabetes,  
 CC rheumatoid arthritis and lupus erythematosus. The products can also be  
 CC used for detection and diagnosis.  
 SQ Sequence 20 AA;

Query Match 88.9%; Score 88; DB 1; Length 20;  
 Best Local Similarity 90.0%; Pred. No. 2.6e-08;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 YRLATRLNERENLRALRY 20  
 ||||| ||||| ||||| |||||  
 Db 1 YRLAIRLNERENLRALRY 20

## RESULT 10

ID R95430 standard; peptide; 20 AA.  
 AC R95430:  
 DT 12-NOV-1996 (first entry)  
 DE HLA-B2702 84-75T/75-84T palindromic.  
 KW HLA; p74; alpha1-helix; human-leucocyte-associated antigen; inhibitor;  
 KW T-cell lysate; membrane protein; mammal; heat shock protein; Hsc70; APC;  
 KW B cell; calcium influx; cytotoxic T lymphocyte; CTL; differentiation;  
 KW cytolysis; antigen presenting cell.  
 OS Synthetic.  
 PN W09513288-A1.  
 PD 18-MAY-1995.  
 PF 10-NOV-1994; U12985.  
 PR 10-NOV-1993; US-150493.  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI Clayberger C, Krensky AM;  
 DR WPI: 95-194027/25.  
 PT Compns. comprising lymphoid surface membrane proteins - which may  
 PT inhibit cytolytic activity and differentiation of CTLs.  
 PS Example; Page 12; 29pp; English.  
 CC R95413, and R95415-R95431 represent palindromes and fragments of  
 CC human-leucocyte-associated antigens. This sequence represents the  
 CC HLA-B2702 84-75T/75-84T palindromic. These sequences can be used to  
 CC isolate the protein p74 from a T-cell lysate. p74 is a T-cell surface  
 CC membrane protein associated with T-cell activation in mammalian T-cells,  
 CC and is also immunologically cross reactive with the heat shock protein  
 CC Hsc70. p74 is found in a limited number of cell types, but is  
 CC particularly expressed on B and T cells. p74 can be isolated by lysis of  
 CC a suitable cell with an amphoteric detergent, and then passed through an  
 CC affinity column containing a covalently bound HLA-B2702 palindromic  
 CC peptide. Compositions comprising the extracellular fragment of p74  
 CC combined with HLA-B2702.60-84 (see R95416), induces calcium influx, and  
 CC inhibits cytotoxic T lymphocyte (CTL) differentiation or cytolysis.  
 CC Candidate compounds can be screened for their effect on the cytolytic  
 CC activity of T-cells, by combining them with the extracellular portion of  
 CC p74 and determining the amount of binding between the candidate compound  
 CC and p74. Modulation of CTL activity can be inhibited in a cellular  
 CC composition containing T-cells and antigen presenting cells (APCs), by  
 CC adding to the mix the extracellular portion of p74, in an amount  
 CC sufficient to compete with p74 for the binding of the p74 ligand.  
 SQ Sequence 20 AA;

Query Match 80.8%; Score 80; DB 1; Length 20;  
 Best Local Similarity 89.5%; Pred. No. 5.2e-07;  
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 YRLATRLNERENLRALRY 19  
 ||||| ||||| ||||| |||||  
 Db 1 YRLAIRLNERENLRALRY 19

## RESULT 11

R48286  
ID R48286 standard; peptide: 25 AA.  
AC R48286:  
DT 15-MAR-1994 (first entry)  
DE Peptide fragment of HLA-B38 antigen.  
KW Human leukocyte antigen; HLA; peptide; transplantation; neoplasia;  
KW parasitic disease; cytotoxic T lymphocyte; modulation.  
OS Synthetic.  
PN WO9317699-A.  
PD 16-SEP-1993.  
PF 25-FEB-1993; U01758.  
PR 02-MAR-1992; US-844716.  
PI (STRD ) UNIV LELAND STANFORD JUNIOR.  
PI Clayberger C, Krensky AM;  
DR WPI: 93-303134/38.  
PT New peptide(s) based on Class I HLA antigen domains - used for  
PT modulating cytotoxic T-lymphocyte activity towards targets  
PS Example 13; Page 39; 61pp; English.  
CC The peptide is used to modulate cytotoxic T-lymphocyte (CTL)  
CC activity, either by inhibition or stimulation. It can be used for  
CC inhibiting CTL toxicity in transplantations, for inducing CTL  
CC activity in parasitic diseases and neoplasia and in studies on viral  
CC infection. The peptide can also be used for identifying CTLs which  
CC bind to it and removing subsets of CTLs from a T-cell composition.  
CC This peptide is derived from the HLA-B38 antigen and corresponds  
CC to the amino acid positions 60-84 of that antigen.  
SQ Sequence 25 AA;

Query Match 50.5%; Score 50; DB 1; Length 25;  
Best Local Similarity 68.8%; Pred. No. 0.051;  
Matches 11; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 TRLNRRERENLRALRY 20  
| | | | | | | | | |  
DB 10 TNTQTYRENLRALRY 25

RESULT 12  
R83093  
ID R83093 standard; peptide: 25 AA.  
AC R83093:  
DT 16-MAY-1996 (first entry)  
DE HLAB38 CTL modulating peptide (B38.6084).  
KW Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;  
KW immunosuppressant; graft versus host disorder; transplantation; therapy;  
KW class I MHC; HLAB38.  
OS Synthetic.  
PN WO9526979-A1.  
PD 12-OCT-1995.  
PF 05-APR-1995; U04349.  
PR 05-APR-1994; US-222851.  
PI (STRD ) UNIV LELAND STANFORD JUNIOR.  
PI Clayberger C, Krensky AM, Parham P;  
DR WPI: 95-358582/46.  
PT Extension of acceptance period of transplants from MHC unmatched  
PT donor hosts - using Class I B75-84 MHC antigen of the recipient  
PT host  
PS Example 13; Page 32; 80pp; English.  
CC R83061-R83085, R83090-R83096 and R92907-R92913 represent fragments of  
CC class I major histocompatibility complex (MHC) antigens. This sequence  
CC corresponds to residues 60-84 of the alpha-1 domain of the class I MHC  
CC HLAB38. These sequences can be used to extend the period of acceptance  
CC by a recipient of a transplant from an MHC unmatched donor. The peptides  
CC are administered to a patient in conjunction with a subtherapeutic amount  
CC of an immunosuppressant. This is administered to the patient for a  
CC limited period of time (compared to the lifetime administration for  
CC current treatments). The peptides particularly modulate (or inhibit) the  
CC activity of the cytotoxic T lymphocytes (CTLs) of the patient.  
SQ Sequence 25 AA;

Query Match 50.5%; Score 50; DB 1; Length 25;  
Best Local Similarity 68.8%; Pred. No. 0.051;  
Matches 11; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 TRLNRRERENLRALRY 20  
| | | | | | | | | |  
DB 10 TNTQTYRENLRALRY 25

Best Local Similarity 68.8%; Pred. No. 0.051;  
Matches 11; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 TRLNRRERENLRALRY 20  
| | | | | | | | | |  
DB 10 TNTQTYRENLRALRY 25

RESULT 13  
R95422  
ID R95422 standard; peptide: 25 AA.  
AC R95422:  
DT 12-NOV-1996 (first entry)  
DE HLAB38.6084.  
KW HLA: p74; alpha-helix; human-leucocyte-associated antigen; inhibitor;  
KW T-cell lysate; membrane protein; mammal; heat shock protein; Hsc70; APC;  
KW B cell; calcium influx; cytotoxic T lymphocyte; CTL; differentiation;  
KW cytolysis; antigen presenting cell.  
OS Synthetic.  
PN WO9513288-A1.  
PD 18-MAY-1995.  
PF 10-NOV-1994; U12985.  
PR 10-NOV-1993; US-150493.  
PI (STRD ) UNIV LELAND STANFORD JUNIOR.  
PI Clayberger C, Krensky AM;  
DR WPI: 95-194027/25.  
PT Compsns. comprising lymphoid surface membrane proteins - which may  
PT inhibit cytolytic activity and differentiation of CTLs.  
PS Example; Page 9; 29pp; English.  
CC R95413, and R95415-R95431 represent palindromes and fragments of  
CC human-leucocyte-associated antigens. These sequences can be used to isolate the protein p74 from  
CC HLAB38.6084. These sequences can be used to isolate the protein p74 from  
CC a T-cell lysate. p74 is a T-cell surface membrane protein associated  
CC with T-cell activation in mammalian T-cells, and is also immunologically  
CC cross reactive with the heat shock protein Hsc70. p74 is found in a  
CC limited number of cell types, but is particularly expressed on B and T  
CC cells. p74 can be isolated by lysis of a suitable cell with an  
CC amphoteric detergent, and then passed through an affinity column  
CC containing a covalently bound HLA-B2702 palindromic peptide.  
CC Compositions comprising the extracellular fragment of p74 combined with  
CC HLA-B2702.60-84 (see R95416), induces calcium influx, and inhibits  
CC cytotoxic T lymphocyte (CTL) differentiation or cytolysis. Candidate  
CC compounds can be screened for their effect on the cytolytic activity of  
CC T-cells, by combining them with the extracellular portion of p74 and  
CC determining the amount of binding between the candidate compound and p74.  
CC Modulation of CTL activity can be inhibited in a cellular composition  
CC containing T-cells and antigen presenting cells (APCs), by adding to the  
CC mix the extracellular portion of p74, in an amount sufficient to compete  
CC with p74 for the binding of the p74 ligand.  
SQ Sequence 25 AA;

Query Match 50.5%; Score 50; DB 1; Length 25;  
Best Local Similarity 68.8%; Pred. No. 0.051;  
Matches 11; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 TRLNRRERENLRALRY 20  
| | | | | | | | | |  
DB 10 TNTQTYRENLRALRY 25

RESULT 14  
Y06801  
ID Y06801 standard; Protein: 184 AA.  
AC Y06801:  
DT 23-JUN-1999 (first entry)  
DE Peptide Seq ID No: 21.  
KW Major histocompatibility complex; MHC; antigen binding clef; allergy;  
KW antigen-specific T-cell; transplant rejection; autoimmune disease;  
KW multiple sclerosis; rheumatoid arthritis; systemic lupus erythematosus;  
KW insulin-dependent diabetes mellitus; immune response; cancer.  
OS Homo sapiens.  
PN WO9914236-A1.

PD 25-MAR-1999.  
 PF 15-SEP-1998; U18244.  
 PR 10-OCT-1997; US-064555.  
 PR 16-SEP-1997; US-064552.  
 PA (BURR/) BURROWS G G.  
 PA (VAND/) VANDENBARK A A.  
 PI Burrows GG, Vandenbark AA;  
 DR WPI: 99-229498/19.  
 PT Recombinant 2-domain major histocompatibility complex molecules used  
 FT to manipulate antigen specific T-cells  
 PS Example 1; Page 67-68; 73pp; English.  
 CC The invention relates to recombinant polypeptides comprising only those  
 CC domains of major histocompatibility complex (MHC) molecules that define  
 CC the antigen binding cleft. These polypeptides can be used to mimic the  
 CC function of mammalian MHC. The recombinant polypeptide comprises  
 CC covalently linked first and second domains, where the first domain is a  
 CC mammalian MHC class II beta1 domain and the second domain is a mammalian  
 CC class I alpha1 domain, or where the first domain is a mammalian MHC  
 CC class I alpha2 domain, and wherein the polypeptide is not a whole MHC class I  
 CC alpha chain. The amino terminus of the second domain is covalently linked  
 CC to the carboxy terminus of the first domain in both cases, optionally  
 CC with a peptide linker sequence. The recombinant MHC polypeptides are  
 CC useful for a wide range of in vitro and in vivo applications, and may be  
 CC used in place of either intact, purified MHC molecules or antigen  
 CC presenting cells that express MHC molecules. The polypeptides can be used  
 CC for the detection, quantification and purification of antigen-specific  
 CC T-cells in biological samples. They can also be used to activate or  
 CC inactivate T-cells and to induce anergy in T-cells. The polypeptides may  
 CC be used for the amelioration of conditions mediated by antigen-specific  
 CC T-cells, e.g. allergies, transplant rejection and autoimmune diseases  
 CC including multiple sclerosis, rheumatoid arthritis, systemic lupus  
 CC erythematosus and insulin-dependent diabetes mellitus. The polypeptides  
 CC may also be used to boost immune responses in certain conditions such as  
 CC cancer and infectious diseases. If a toxic molecule is attached to the  
 CC polypeptides, then they can also be used to kill specific T-cells.  
 SQ Sequence 184 AA;

Query Match 50.5%; Score 50; DB 1; Length 184;  
 Best Local Similarity 68.8%; Pred. No. 0.46;  
 Matches 11; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 5 TRLNERENLRALRY 20  
 | | | | | | | | | |  
 Db 69 TNTQTYRENLRALRY 84

RESULT 15  
 R03142  
 ID R03142 standard; protein; 362 AA.  
 AC R03142; 1991 (first entry)  
 DT 19-MAR-1991 (first entry)  
 DE Sequence of HLA-B\*52 antigen.  
 KW Probe: HLA class I DNA; immunogen.  
 OS Homo sapiens.  
 PN EP-354580-A.  
 PD 14-FEB-1990.  
 PF 11-AUG-1989.  
 PR 11-AUG-1988; JP-200758.  
 PA (OLYU) Olympus Optical Co., Ltd.  
 PI Kano K, Takiguchi;  
 DR WPI: 90-046289/07.  
 PT New DNA for class I human leucocyte antigens and derived probes and  
 FT transformed cells, useful for DNA typing, as immunogens etc.  
 PS Disclosure; Page 13; 23pp; English.  
 CC The HLA class I DNA can be used as a source of probes for use in DNA  
 CC typing. Transformed cells, which are useful as immunogens, can be  
 CC obtained by introducing these DNAs into eucaryotic cells.  
 SQ Sequence 362 AA;

Query Match 50.5%; Score 50; DB 1; Length 362;

Best Local Similarity 68.8%; Pred. No. 0.98;  
 Matches 11; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 5 TRLNERENLRALRY 20  
 | | | | | | | | | |  
 Db 93 TNTQTYRENLRALRY 108

Search completed: February 8, 2000, 04:05:43  
 Job time: 9360 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 7, 2000, 18:04:38 ; Search time 111.22 seconds  
(without alignments)  
8.482 Million cell updates/sec

Title: US-08-653-294-28

Perfect score: 99

Sequence: 1 YRLATRLNERENLRALRY 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 142080 seqs, 47169319 residues

Total number of hits satisfying chosen parameters: 142080

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : PIR\_62.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50	50.5	273	I38509	MHC class I histoc
2	50	50.5	274	I54463	MHC HLA-B38 chain
3	50	50.5	362	B30345	MHC class I histoc
4	50	50.5	362	A45834	MHC class I histoc
5	50	50.5	362	I84486	transmembrane glyco
6	50	50.5	362	A30345	MHC class I histoc
7	50	50.5	362	I59633	MHC HLA-B transmem
8	50	50.5	362	S24434	class I histoc
9	50	50.5	362	I37120	MHC class I histoc
10	49	49.5	354	I59308	class I histoc
11	49	49.5	354	I80168	class I histoc
12	49	49.5	354	I80167	class I histoc
13	49	49.5	355	I80169	class I histoc
14	49	49.5	355	I80171	class I histoc
15	49	49.5	359	IHLH12	MHC class I histoc
16	49	49.5	362	IHLH08	MHC class I histoc
17	49	49.5	362	JH0341	class I histoc
18	49	49.5	362	JH0539	class I histoc
19	49	49.5	362	JH0540	class I histoc
20	49	49.5	362	I62045	gene HLA B-1517 pr
21	49	49.5	362	I84490	lymphocyte antigen
22	49	49.5	362	I37521	HLA-Bw57.2 antigen
23	49	49.5	363	S07113	class I histoc
24	49	49.5	363	S03537	class I histoc
25	49	49.5	364	D35997	MHC class I histoc
26	49	49.5	365	S77963	MHC class I histoc
27	49	49.5	365	I54416	HLA-AW24 protein -
28	49	49.5	365	I54493	MHC class I histoc
29	45	45.5	355	I37516	HLA-B alpha-chain
30	45	45.5	362	S25415	class I histoc

31	45	45.5	362	2	A45850	MHC class I histoc
32	45	45.5	362	2	I61861	MHC HLA-B44.2 chai
33	45	45.5	362	2	I54442	MHC class I histoc
34	44	44.4	274	1	HLH032	MHC class I histoc
35	44	44.4	364	2	A35997	MHC class I histoc
36	44	44.4	445	2	S67147	hypothetical prote
37	43	43.4	455	2	F64617	hypothetical prote
38	43	43.4	2013	2	C71610	probable membrane
39	42	42.4	137	2	I80174	class I histoc
40	42	42.4	425	1	FOMVGC	gag polyprotein -
41	42	42.4	536	1	FOMVMD	gag polyprotein -
42	42	42.4	790	1	FOMVHZ	gag-kit polyprotei
43	42	42.4	912	2	F71433	probable growth re
44	42	42.4	1110	1	A70652	probable serine/th
45	42	42.4	1228	2	I40468	surface layer prot

#### ALIGNMENTS

RESULT 1

I38509

MHC class I histocompatibility antigen - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 06-Sep-1996 #sequence\_revision 06-Sep-1996 #text\_change 23-Jul-1999

C:Accession: I38509

R:Cereb. N.; Choi, J.W.; Riu, K.Z.; Yang, S.Y.

Tissue Antigens 44, 271-273, 1994

A:Title: HLA-B\*5105, a newly identified B51 IEF variant.

A:Reference number: I38509; MUID:95176331

A:Accession: I38509

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-273 <RES>

A:Cross-references: EMBL:U06697; NID:g469544; PIDN:AAA92997.1; PID:g469545

C:Genetics:

A:Gene: GDB:HLA-B

A:Cross-references: GDB:I20048; OMIM:142830

A:Map position: 6p21.3-6p21.3

C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 50.5%; Score 50; DB 2; Length 273;

Best Local Similarity 68.8%; Pred. No. 1;

Matches 11; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 TRLNERENLRALRY 20

Db 68 TTTQTYRENLRLRY 83

RESULT 2

I54463

MHC HLA-B38 chain - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 07-Jun-1996 #sequence\_revision 07-Jun-1996 #text\_change 23-Jul-1999

C:Accession: I54463

R:Mueller, C.A.; Engler-Blum, G.; Gekeler, V.; Steiert, I.; Weiss, E.; Schmidt, H.

Immunogenetics 30, 200-207, 1989

A:Title: Genetic and serological heterogeneity of the supertypic HLA-B locus specific

A:Reference number: I54463; MUID:89379286

A:Accession: I54463

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-274 <RES>

A:Cross-references: GB:M29864; NID:g187674; PIDN:AAA36222.1; PID:g187675

C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 50.5%; Score 50; DB 2; Length 274;

Best Local Similarity 68.8%; Pred. No. 1.1;

Matches 11; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 TRLNERENLRALRY 20  
| | | | | | | | | |  
DB 69 TNTQTYRENLRALRY 84

## RESULT 3

MHC class I histocompatibility antigen HLA-Bw52 precursor - human  
C:Species: Homo sapiens (man)  
C:Date: 29-Jan-1990 #sequence\_revision 29-Jan-1990 #text\_change 16-Feb-1997  
C:Accession: B30345  
R:Hayashi, H.; Ennis, P.D.; Ariga, H.; Salter, R.D.; Parham, P.; Kano, K.; Takiguchi, M.  
J. Immunol. 142, 306-311, 1989  
A:Title: HLA-B51 and HLA-Bw52 differ by only two amino acids which are in the helical re  
A:Reference number: A30345; MUID:89080265  
A:Accession: B30345  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-362 <HAY>  
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology  
C:Keywords: transmembrane protein  
F:220-285/Domain: immunoglobulin homology <IMM>

Query Match 50.5%; Score 50; DB 2; Length 362;  
Best Local Similarity 68.8%; Pred. No. 1.4;  
Matches 11; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 TRLNERENLRALRY 20  
| | | | | | | | | |  
DB 93 TNTQTYRENLRALRY 108

## RESULT 4

MHC class I histocompatibility antigen HLA-B53 alpha chain precursor - human  
C:Species: Homo sapiens (man)  
C:Date: 03-Jun-1993 #sequence\_revision 03-Jun-1993 #text\_change 23-Jul-1999  
C:Accession: A45834  
R:Hayashi, H.; Ooba, T.; Nakayama, S.; Sekimata, M.; Kano, K.; Takiguchi, M.  
Immunogenetics 32, 195-199, 1990  
A:Title: Allospecificities between HLA-Bw53 and HLA-B35 are generated by substitution of  
A:Reference number: A45834; MUID:91033941  
A:Accession: A45834  
A:Molecule type: DNA  
A:Residues: 1-362 <HAY>  
A:Cross-references: GB:M58636; NID:g187756; PIDN:AAA36228.1; PID:g187757; GB:M33574  
A:Note: this allele is designated B\*5301  
C:Genetics:

A:Gene: GDB:HLA-B  
A:Cross-references: GDB:120048; OMIM:142830  
A:Map position: 6p21.3-6p21.3  
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology  
C:Keywords: glycoprotein; heterodimer; transmembrane protein  
F:1-24/Domain: signal sequence #status predicted <SIG>  
F:220-285/Domain: immunoglobulin homology <IMM>  
F:110/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 50.5%; Score 50; DB 2; Length 362;  
Best Local Similarity 68.8%; Pred. No. 1.4;  
Matches 11; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 TRLNERENLRALRY 20  
| | | | | | | | | |  
DB 93 TNTQTYRENLRALRY 108

## RESULT 5

MHC class I histocompatibility antigen HLA-B51 precursor - human  
C:Species: Homo sapiens (man)  
C:Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 23-Jul-1999  
C:Accession: A45834  
R:Hayashi, H.; Ooba, T.; Nakayama, S.; Sekimata, M.; Kano, K.; Takiguchi, M.  
Immunogenetics 32, 195-199, 1990  
A:Title: Allospecificities between HLA-Bw53 and HLA-B35 are generated by substitution of  
A:Reference number: A45834; MUID:91033941  
A:Accession: A45834  
A:Molecule type: DNA  
A:Residues: 1-362 <HAY>  
A:Cross-references: GB:M58636; NID:g187756; PIDN:AAA36228.1; PID:g187757; GB:M33574  
A:Note: this allele is designated B\*5301  
C:Genetics:

A:Gene: GDB:HLA-B  
A:Cross-references: GDB:120048; OMIM:142830  
A:Map position: 6p21.3-6p21.3  
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology  
C:Keywords: glycoprotein; heterodimer; transmembrane protein  
F:1-24/Domain: signal sequence #status predicted <SIG>  
F:220-285/Domain: immunoglobulin homology <IMM>  
F:110/Binding site: carbohydrate (Asn) (covalent) #status predicted

C:Accession: I84486  
R:Hildebrand, W.H.; Domena, J.D.; Shen, S.Y.; Lau, M.; Terasaki, P.I.; Bunce, M.; Mar  
Tissue Antigens 43, 209-218, 1994  
A:Title: HLA-B15: a widespread and diverse family of HLA-B alleles.  
A:Reference number: I38421; MUID:94367483  
A:Accession: I84486  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-362 <RES>  
A:Cross-references: GB:L15005; NID:g493154; PIDN:AAA56832.1; PID:g493155  
C:Genetics:  
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology  
C:Keywords: glycoprotein

Query Match 50.5%; Score 50; DB 2; Length 362;  
Best Local Similarity 68.8%; Pred. No. 1.4;  
Matches 11; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 TRLNERENLRALRY 20  
| | | | | | | | | |  
DB 93 TNTQTYRENLRALRY 108

## RESULT 6

MHC class I histocompatibility antigen HLA-B51 precursor - human  
C:Species: Homo sapiens (man)  
C:Date: 29-Jan-1990 #sequence\_revision 29-Jan-1990 #text\_change 23-Jul-1999  
C:Accession: A30345; I68746; I37499  
R:Hayashi, H.; Ennis, P.D.; Ariga, H.; Salter, R.D.; Parham, P.; Kano, K.; Takiguchi, M.  
J. Immunol. 142, 306-311, 1989  
A:Title: HLA-B51 and HLA-Bw52 differ by only two amino acids which are in the helical  
A:Reference number: A30345; MUID:89080265  
A:Accession: A30345  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-362 <HAY>  
R:Pohla, H.; Kuon, W.; Tabaczewski, P.; Doerner, C.; Weiss, E.H.  
Immunogenetics 29, 297-307, 1989  
A:Title: Allelic variation in HLA-B and HLA-C sequences and the evolution of the HLA-  
A:Reference number: I54457; MUID:89233295  
A:Accession: I68746  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-362 <RES>  
A:Cross-references: GB:L41087; NID:g735900; PIDN:AAA64513.1; PID:g735902  
R:Steinle, A.; Schendel, D.J.  
Tissue Antigens 44, 268-270, 1994  
A:Title: HLA class I alleles of LCL 721 and 174 x CEM.T2 (T2).  
A:Reference number: I37499; MUID:95176330  
A:Accession: I37499  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-206 <RE2>  
A:Cross-references: EMBL:Z46808; NID:g599783; PIDN:CAA86838.1; PID:g599784  
C:Genetics:

A:Gene: GDB:HLA-B  
A:Cross-references: GDB:120048; OMIM:142830  
A:Map position: 6p21.3-6p21.3  
A:Introns: 25/1; 115/1; 207/1; 299/1; 338/1; 349/1  
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology  
C:Keywords: transmembrane protein  
F:220-285/Domain: immunoglobulin homology <IMM>

Query Match 50.5%; Score 50; DB 2; Length 362;  
Best Local Similarity 68.8%; Pred. No. 1.4;  
Matches 11; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 TRLNERENLRALRY 20  
| | | | | | | | | |  
DB 93 TNTQTYRENLRALRY 108

Db 93 TNTQTYRENRLRY 108

## RESULT 7

I59633  
MHC HLA-B transmembrane glycoprotein - human

C:Species: Homo sapiens (man)  
C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 23-Jul-1999  
C:Accession: I59633  
R:Hildebrand, W.H.; Domena, J.D.; Parham, P.  
Tissue Antigens 41, 190-195, 1993  
A:Title: Primary structure shows HLA-B59 to be a hybrid of HLA-B55 and HLA-B51, and not  
A:Reference number: I59633; MUID:93369833  
A:Accession: I59633  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-362 <RES>  
A:Cross-references: GB:I07743; NID:G388314; PIDN:AAA59621.1; PID:G388315  
C:Genetics:  
A:Gene: GDB:HLA-B  
A:Cross-references: GDB:I20048; OMIM:142830  
A:Map position: 6p21.3-6p21.3  
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology  
C:Keywords: glycoprotein

Query Match 50.5%; Score 50; DB 2; Length 362;  
Best Local Similarity 68.8%; Pred. No. 1.4;  
Matches 11; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 TRLNERENRLRY 20

Db 93 TNTQTYRENRLRY 108

## RESULT 8

S24434

class I histocompatibility antigen HLA-B-52012 precursor - human

C:Species: Homo sapiens (man)  
C:Date: 22-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 07-Nov-1997  
C:Accession: S24434  
R:Watkins, D.I.; McAdam, S.N.; Liu, X.; Strang, C.R.; Milford, E.L.; Levine, C.G.; Garber, R.; McAdam, S.N.; Boyson, J.E.; Liu, X.; Garber, T.L.; Hughes, A.L.; Bontrorp, R.E.; Wat  
Nature 357, 329-333, 1992  
A:Title: New recombinant HLA-B alleles in a tribe of South American Amerindians indicate  
A:Reference number: S24027; MUID:92269956  
A:Accession: S24434  
A:Molecule type: mRNA  
A:Residues: 1-362 <WAT>  
C:Genetics:  
A:Gene: HLA-B-52012  
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology  
C:Keywords: glycoprotein; transmembrane protein  
F:1-24/Domain: signal sequence #status predicted <SIG>  
F:25-307/Product: class I histocompatibility antigen HLA-B-52012 #status predicted <MAT>  
F:25-307/Domain: extracellular #status predicted <EXT>  
F:115-206/Domain: alpha-2 #status predicted <EX2>  
F:220-285/Domain: immunoglobulin homology <IMM>  
F:308-331/Domain: transmembrane #status predicted <TMW>  
F:332-362/Domain: intracellular #status predicted <INT>  
F:110/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:125-188,227-283/Disulfide bonds: #status predicted

Query Match 50.5%; Score 50; DB 2; Length 362;  
Best Local Similarity 68.8%; Pred. No. 1.4;  
Matches 11; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 TRLNERENRLRY 20

Db 93 TNTQTYRENRLRY 108

## RESULT 9

I37120

MHC class I histocompatibility antigen HLA-B51 alpha chain precursor - human (isolate  
C:Species: Homo sapiens (man)  
A:Variety: isolate B\*5104  
C:Date: 12-Aug-1996 #sequence\_revision 12-Aug-1996 #text\_change 23-Jul-1999  
R:Belich, M.P.; Madrigal, J.A.; Hildebrand, W.H.; Zemmour, J.; Williams, R.C.; Luz, R  
Nature 357, 326-329, 1992  
A:Title: Unusual HLA-B alleles in two tribes of Brazilian Indians.  
A:Reference number: I37120; MUID:92269955

A:Accession: I37120  
A:Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-362 <RES>

A:Cross-references: EMBL:Z15143; NID:G28234; PIDN:CAA78849.1; PID:G28235  
A:Experimental source: cell line KRC 005; isolate B\*5104  
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

C:Keywords: transmembrane protein

F:1-24/Domain: signal sequence #status predicted <SIG>

F:25-362/Product: MHC class I histocompatibility antigen HLA-B51 #status predicted <M

F:220-285/Domain: immunoglobulin homology <IMM>

Query Match 50.5%; Score 50; DB 2; Length 362;  
Best Local Similarity 68.8%; Pred. No. 1.4;  
Matches 11; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 TRLNERENRLRY 20

Db 93 TNTQTYRENRLRY 108

## RESULT 10

I59308

class I histocompatibility antigen - pygmy chimpanzee (fragment)

C:Species: Pan paniscus (Pygmy chimpanzee, bonobo)  
C:Date: 31-May-1996 #sequence\_revision 31-May-1996 #text\_change 23-Jul-1999  
C:Accession: I59308  
R:McAdam, S.N.; Boyson, J.E.; Liu, X.; Garber, T.L.; Hughes, A.L.; Bontrorp, R.E.; Wat

Proc. Natl. Acad. Sci. U.S.A. 91, 5893-5897, 1994

A:Title: A uniquely high level of recombination at the HLA-B locus.

A:Reference number: I59308; MUID:94286544

A:Accession: I59308

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-354 <RES>

A:Cross-references: EMBL:U05575; NID:G454767; PIDN:AAA50178.1; PID:G454768

C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 49.5%; Score 49; DB 2; Length 354;  
Best Local Similarity 100.0%; Pred. No. 2;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RENRLRY 20

Db 91 RENRLRY 100

## RESULT 11

I80168

class I histocompatibility antigen - chimpanzee (fragment)

C:Species: Pan troglodytes (chimpanzee)

C:Date: 24-May-1996 #sequence\_revision 24-May-1996 #text\_change 23-Jul-1999

C:Accession: I80168

R:McAdam, S.N.; Boyson, J.E.; Liu, X.; Garber, T.L.; Hughes, A.L.; Bontrorp, R.E.; Wat

Proc. Natl. Acad. Sci. U.S.A. 91, 5893-5897, 1994

A:Title: A uniquely high level of recombination at the HLA-B locus.

A:Reference number: I59308; MUID:94286544

A:Accession: I80168

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-354 <RES>

A:Cross-references: EMBL:U05579; NID:g454775; PIDN:AAA50182.1; PID:g454776  
 C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 49.5%; Score 49; DB 2; Length 354;  
 Best Local Similarity 100.0%; Pred. No. 2;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 RENLIRALRY 20  
 |||||  
 Db 91 RENLIRALRY 100

## RESULT 12

180167 class I histocompatibility antigen - pygmy chimpanzee (fragment)

C:Species: Pan paniscus (pygmy chimpanzee, bonobo)  
 C:Date: 24-May-1996 #sequence\_revision 24-May-1996 #text\_change 23-Jul-1999  
 C:Accession: I80167

R:McAdam, S.N.; Boyson, J.E.; Liu, X.; Garber, T.L.; Hughes, A.L.; Bontrop, R.E.; Watkins  
 Proc. Natl. Acad. Sci. U.S.A. 91, 5893-5897, 1994

A:Title: A uniquely high level of recombination at the HLA-B locus.

A:Reference number: I59308; MUID:94286544

A:Accession: I80167

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-354 <RES>

A:Cross-references: EMBL:U05578; NID:g454773; PIDN:AAA50181.1; PID:g454774  
 C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 49.5%; Score 49; DB 2; Length 354;  
 Best Local Similarity 100.0%; Pred. No. 2;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 RENLIRALRY 20  
 |||||  
 Db 91 RENLIRALRY 100

## RESULT 13

180169 class I histocompatibility antigen - chimpanzee (fragment)

C:Species: Pan troglodytes (chimpanzee)  
 C:Date: 24-May-1996 #sequence\_revision 24-May-1996 #text\_change 23-Jul-1999  
 C:Accession: I80169

R:McAdam, S.N.; Boyson, J.E.; Liu, X.; Garber, T.L.; Hughes, A.L.; Bontrop, R.E.; Watkins  
 Proc. Natl. Acad. Sci. U.S.A. 91, 5893-5897, 1994

A:Title: A uniquely high level of recombination at the HLA-B locus.

A:Reference number: I59308; MUID:94286544

A:Accession: I80169

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-355 <RES>

A:Cross-references: EMBL:U05580; NID:g454777; PIDN:AAA50183.1; PID:g454778  
 C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 49.5%; Score 49; DB 2; Length 355;  
 Best Local Similarity 100.0%; Pred. No. 2;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 RENLIRALRY 20  
 |||||  
 Db 91 RENLIRALRY 100

## RESULT 14

180171 class I histocompatibility antigen - chimpanzee (fragment)

C:Species: Pan troglodytes (chimpanzee)  
 C:Date: 24-May-1996 #sequence\_revision 24-May-1996 #text\_change 23-Jul-1999  
 C:Accession: I80171

R:McAdam, S.N.; Boyson, J.E.; Liu, X.; Garber, T.L.; Hughes, A.L.; Bontrop, R.E.; Wat  
 Proc. Natl. Acad. Sci. U.S.A. 91, 5893-5897, 1994  
 A:Title: A uniquely high level of recombination at the HLA-B locus.

A:Reference number: I59308; MUID:94286544

A:Accession: I80171

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-355 <RES>

A:Cross-references: EMBL:U05582; NID:g454781; PIDN:AAA50185.1; PID:g454782  
 C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 49.5%; Score 49; DB 2; Length 355;  
 Best Local Similarity 100.0%; Pred. No. 2;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 RENLIRALRY 20  
 |||||  
 Db 91 RENLIRALRY 100

## RESULT 15

HLH012

MHC class I histocompatibility antigen HLA alpha chain precursor (clone PHLA 12.4) -

C:Species: Homo sapiens (man)

C:Date: 05-Apr-1983 #sequence\_revision 05-Apr-1983 #text\_change 22-Jun-1999

C:Accession: A02189

R:Malissen, M.; Malissen, B.; Jordan, B.R.

Proc. Natl. Acad. Sci. U.S.A. 79, 893-897, 1982

A:Title: Exon/intron organization and complete nucleotide sequence of an HLA gene.

A:Reference number: A02189; MUID:82151002

A:Accession: A02189

A:Molecule type: DNA

A:Residues: 1-359 <MAL>

A:Cross-references: GB:J00191; GB:V00526; NID:gl87600; PIDN:AAA36218.1; PID:g386873  
 C:Comment: The seven exons correspond approximately to the domain structure of this c

C:Genetics:

A:Map position: 6p21.3

A:Introns: 22/1; 112/1; 204/1; 296/1; 335/1; 345/1

C:Superfamily: class I histocompatibility antigen; immunoglobulin homology  
 C:Keywords: duplication; glycoprotein; heterodimer; transmembrane protein; transplant

F:1-21/Domain: signal sequence #status predicted <SIG>  
 F:22-359/Product: class I histocompatibility antigen HLA alpha chain #status predicted

F:22-304/Domain: extracellular #status predicted <EXT>

F:22-111/Domain: alpha-1 <EX1>

F:112-203/Domain: alpha-2 <EX2>

F:112-203/Domain: immunoglobulin homology <IMM>

F:305-329/Domain: transmembrane #status predicted <TM>

F:335-359/Domain: intracellular #status predicted <INT>

F:107/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:224-280/Disulfide bonds: #status predicted

Query Match 49.5%; Score 49; DB 1; Length 359;  
 Best Local Similarity 100.0%; Pred. No. 2;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 RENLIRALRY 20  
 |||||  
 Db 96 RENLIRALRY 105

Search completed: February 7, 2000, 18:04:38  
 Job time: 22204 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2000, 01:25:59 ; Search time 68.63 seconds  
(without alignments)  
8.703 Million cell updates/sec

Title: US-08-653-294-28

Perfect score: 99

Sequence: 1 YRLATRLNERENLRALRY 20

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 82229 seqs, 29864866 residues

Total number of hits satisfying chosen parameters: 82229

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : SwissProt\_38.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	50	50.5	362	1	1B47_HUMAN	P30487 homo sapien
2	50	50.5	362	1	1B49_HUMAN	P18464 homo sapien
3	50	50.5	362	1	1B52_HUMAN	P30489 homo sapien
4	50	50.5	362	1	1B53_HUMAN	P30490 homo sapien
5	50	50.5	362	1	1B54_HUMAN	P30491 homo sapien
6	49	49.5	359	1	1B01_PANTR	P13750 pan troglod
7	49	49.5	362	1	1B01_GORGO	P30379 gorilla gor
8	49	49.5	362	1	1B02_GORGO	P30380 gorilla gor
9	49	49.5	362	1	1B03_GORGO	P30381 gorilla gor
10	49	49.5	362	1	1B15_HUMAN	P10317 homo sapien
11	49	49.5	362	1	1B60_HUMAN	P18465 homo sapien
12	49	49.5	362	1	1B61_HUMAN	P30497 homo sapien
13	49	49.5	362	1	1B62_HUMAN	P10319 homo sapien
14	49	49.5	362	1	1B62_HUMAN	P01893 homo sapien
15	49	49.5	365	1	1A23_HUMAN	P30447 homo sapien
16	49	49.5	365	1	1A24_HUMAN	P05534 homo sapien
17	45	45.5	362	1	1B05_HUMAN	P30461 homo sapien
18	45	45.5	362	1	1B41_HUMAN	P30481 homo sapien
19	45	45.5	362	1	1B42_HUMAN	P30482 homo sapien
20	44	44.4	365	1	1A25_HUMAN	P18462 homo sapien
21	44	44.4	365	1	1A32_HUMAN	P10314 homo sapien
22	42	42.4	414	1	GAG_FSVHZ	P04322 feline sarc
23	42	42.4	425	1	GAG_FSVGA	P03337 feline sarc
24	42	42.4	536	1	GAG_FSVMD	P03340 feline sarc
25	42	42.4	580	1	GAG_FLV	P10262 feline leuk
26	42	42.4	1228	1	SLAP_BACST	P35825 bacillus st
27	42	42.4	1807	1	TSC2_HUMAN	P49815 homo sapien
28	42	42.4	1809	1	TSC2_RAT	P49816 rattus norv
29	41	41.4	162	1	PUR6_BACSV	P12044 bacillus su
30	41	41.4	2077	1	TEGU_HSV6G	P30002 herpes simp
31	41	41.4	2077	1	TEGU_HSV6U	P52340 herpes simp
32	40.5	40.9	837	1	YLE5_CAEEL	P46941 caenorhabdi
33	40	40.4	277	1	YDIA_ECOLI	P03822 escherichia
34	40	40.4	359	1	1B40_HUMAN	P10320 homo sapien

35 40 40.4 924 1 ORC1\_DROME O16810 drosophila  
36 39 39.4 234 1 YGGS\_ECOLI P52054 escherichia  
37 39 39.4 354 1 GRA3\_USTHO O14438 ustilago ho  
38 39 39.4 365 1 1A04\_GORGO P03078 gorilla gor  
39 39 39.4 400 1 NUCM\_PROWI Q37619 protocheca  
40 39 39.4 432 1 MDHD\_SORVU P37229 sorghum vul  
41 39 39.4 454 1 BLH1\_YEAST Q01532 saccharomyc  
42 39 39.4 506 1 MKK2\_YEAST P32491 saccharomyc  
43 39 39.4 518 1 G6PD\_DROYA O27638 drosophila  
44 39 39.4 591 1 GLMS\_AQUAE O65648 a glucosami  
45 39 39.4 856 1 PFSA\_AQUAE O67899 aquifex aeo

## ALIGNMENTS

RESULT 1  
ID 1B47\_HUMAN STANDARD; PRT; 362 AA.  
AC P30487;  
DT 01-APR-1993 (Rel. 25, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 01-FEB-1996 (Rel. 33, Last annotation update)  
DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-49(B-21) B\*4901 ALPHA CHAIN  
DE PRECURSOR.  
GN HLA-B OR HLAB.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Homnidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 89235215.  
RA PARHAM P., LAWLOD D.A., LOMEN C.E., ENNIS P.D.;  
RT "Diversity and diversification of HLA-A,B,C alleles."  
RL J. Immunol. 142:3937-3950(1989).  
RN [2]  
RP REVISION TO 78.  
RX MEDLINE; 93056529.  
RA HILDEBRAND W.H., MADRIGAL J.A., BELICH M.P., ZEMMOUR J., WARD F.E.,  
RT "Serologic cross-reactivities poorly reflect allelic relationships in the HLA-B\*2 and HLA-B\*21 groups. Dominant epitopes of the alpha 2 helix."  
RL J. Immunol. 149:3563-3568(1992).  
CC - FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE IMMUNE SYSTEM.  
CC - SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-MICROGLOBULIN).  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).  
CC EMBL; M24037; AAA02950.1; ..  
CC HSP; P30491; 1A1M.  
CC MIM; 142830; ..  
CC PROSITE; PS00290; IG\_MHC; 1.  
CC PFAM; PF00047; Ig; 1.  
CC PFAM; PF00129; MHC.I; 1.  
CC MHC I; Transmembrane; Glycoprotein; Signal.  
CC SIGNAL 1 24  
CC CHAIN 25 362  
CC HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-49(B-21) B\*4901 ALPHA CHAIN.  
CC EXTRACELLULAR ALPHA-1.  
CC EXTRACELLULAR ALPHA-2.  
CC EXTRACELLULAR ALPHA-3.  
CC CONNECTING PEPTIDE.  
CC CYTOPLASMIC TAIL.

FT CARBOHYD 110 110 BY SIMILARITY.  
 FT FT DISULFID 125 188 BY SIMILARITY.  
 FT FT DISULFID 227 283 BY SIMILARITY.  
 SQ SEQUENCE 362 AA; 40581 MW; E99F82F CRC32;

Query Match 50.5%; Score 50; DB 1; Length 362;  
 Best Local Similarity 68.8%; Pred. No. 0.46;  
 Matches 11; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 TRLNERENLRALRY 20  
 DB 93 TQTQTYRENLRALRY 108

RESULT 2  
 ID 1B49\_HUMAN STANDARD; PRT; 362 AA.  
 AC P18464;  
 DT 01-NOV-1990 (Rel. 16, Created)  
 DT 01-NOV-1990 (Rel. 16, Last sequence update)  
 DT 01-FEB-1996 (Rel. 33, Last annotation update)  
 DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-51(B-5) B\*5101 ALPHA CHAIN  
 DE PRECURSOR.  
 GN HLA-B OR HLAB.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 90207291.  
 RA ENNIS P.D., ZEMMOUR J., SALTER R.D., PARHAM P.;  
 RT "Rapid cloning of HLA-A,B CDNA by using the polymerase chain  
 reaction: frequency and nature of errors produced in amplification.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 87:2833-2837(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 89080265.  
 RA HAYASHI H., ENNIS P.D., ARIGA H., SALTER R.D., PARHAM P., KANO K.,  
 RA TAKIGUCHI M.;  
 RT "HLA-B51 and HLA-Bw52 differ by only two amino acids, which are in the  
 helical region of the alpha 1 domain";  
 RL J. Immunol. 142:306-311(1989).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 89233295.  
 RA POHLA H., KUON W., TABACZEWSKI P., DOERNER C., WEISS E.H.;  
 RT "Allelic variation in HLA-B and HLA-C sequences and the evolution of  
 the HLA-B alleles";  
 RL Immunogenetics 29:297-307(1989).  
 CC -1- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO  
 CC THE IMMUNE SYSTEM.  
 CC -1- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
 CC MICROGLOBULIN).

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

EMBL; M32319; AAA36232.1;  
 DR EMBL; M22792; AAA59620.1; ALT\_SEQ.  
 DR EMBL; M22786; AAA59620.1; JOINED.  
 DR EMBL; M22787; AAA59620.1; JOINED.  
 DR EMBL; M22788; AAA59620.1; JOINED.  
 DR EMBL; M22789; AAA59620.1; JOINED.  
 DR EMBL; M22790; AAA59620.1; JOINED.  
 DR EMBL; M22791; AAA59620.1; JOINED.  
 DR EMBL; L41087; AAA64513.1;  
 DR EMBL; L41086; AAA64513.1; JOINED.

DR PIR; A30345; A30345.  
 DR PIR; A30548; A30548.  
 DR HSSP; P30491; 1A1M.  
 DR MIM; 142830;  
 DR PROSITE; PS00290; IG\_MHC; 1.  
 DR PFAM; PF00047; Ig; 1.  
 DR PFAM; PF00129; MHC\_I; 1.  
 KW MHC I; Transmembrane; Glycoprotein; Signal.  
 FT SIGNAL 1 24  
 FT CHAIN 25 362 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,  
 FT B-51(B-5) B\*5101 ALPHA CHAIN.  
 FT DOMAIN 25 114 EXTRACELLULAR ALPHA-1.  
 FT DOMAIN 115 206 EXTRACELLULAR ALPHA-2.  
 FT DOMAIN 207 298 EXTRACELLULAR ALPHA-3.  
 FT DOMAIN 299 308 CONNECTING PEPTIDE.  
 FT TRANSMEM 309 332  
 FT DOMAIN 333 362 CYTOPLASMIC TAIL.  
 FT CARBOHYD 110 110 BY SIMILARITY.  
 FT DISULFID 125 188 BY SIMILARITY.  
 FT DISULFID 227 283 BY SIMILARITY.  
 SQ SEQUENCE 362 AA; 40566 MW; 4D846F30 CRC32;

Query Match 50.5%; Score 50; DB 1; Length 362;  
 Best Local Similarity 68.8%; Pred. No. 0.46;  
 Matches 11; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 TRLNERENLRALRY 20  
 DB 93 TQTQTYRENLRALRY 108

RESULT 3  
 ID 1B52\_HUMAN STANDARD; PRT; 362 AA.  
 AC P30489;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 01-APR-1993 (Rel. 25, Last annotation update)  
 DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-51(B-5) B\*5104 ALPHA CHAIN  
 DE PRECURSOR.  
 GN HLA-B OR HLAB.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 92269955.  
 RA BELICH M.P., MADRIGAL J.A., HILDEBRAND W.H., ZEMMOUR J.,  
 RA WILLIAMS R.C., LUZ R., PETZL-ERLER M.L., PARHAM P.;  
 RT "Unusual HLA-B alleles in two tribes of Brazilian Indians.";  
 RL Nature 357:326-329(1992).  
 CC -1- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO  
 CC THE IMMUNE SYSTEM.  
 CC -1- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
 CC MICROGLOBULIN).

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

EMBL; 215143; CAAT78849.1;  
 DR HSSP; P30491; 1A1M.  
 DR MIM; 142830;  
 DR PROSITE; PS00290; IG\_MHC; 1.  
 DR PFAM; PF00047; Ig; 1.  
 DR PFAM; PF00129; MHC\_I; 1.  
 KW MHC I; Transmembrane; Glycoprotein; Signal.  
 FT SIGNAL 1 24

FT CHAIN 25 362 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,  
B\*51(B-5) B\*5104 ALPHA CHAIN.  
FT DOMAIN 25 114 EXTRACELLULAR ALPHA-1.  
FT DOMAIN 115 206 EXTRACELLULAR ALPHA-2.  
FT DOMAIN 207 298 EXTRACELLULAR ALPHA-3.  
FT DOMAIN 299 308 CONNECTING PEPTIDE.  
FT TRANSMEM 309 332  
FT DOMAIN 333 362 CYTOPLASMIC TAIL.  
FT CARBOHYD 110 110 BY SIMILARITY.  
FT DISULFID 125 188 BY SIMILARITY.  
FT DISULFID 227 283 BY SIMILARITY.  
SQ SEQUENCE 362 AA; 40560 MW; F22F08AB CRC32;

Query Match 50.5%; Score 50; DB 1; Length 362;  
Best Local Similarity 68.8%; Pred. No. 0.46;  
Matches 11; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 5 TRLNERENLRALRY 20  
I | | | | | | | | | |  
Db 93 TMTQTYRENLRALRY 108

RESULT 4  
1B53\_HUMAN  
ID 1B53\_HUMAN STANDARD; PRT; 362 AA.  
AC P30490;  
DT 01-APR-1993 (Rel. 25, Created)  
DT 01-APR-1993 (Rel. 25, Last sequence update)  
DT 01-APR-1993 (Rel. 25, Last annotation update)  
DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, BW-52(B-5) B\*5201 ALPHA CHAIN  
DE PRECURSOR.  
GN HLA-B OR HLAB.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Homnidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 89080265.  
RA HAYASHI H., ENNIS P.D., ARIGA H., SALTER R.D., PARHAM P., KANO K.,  
TAKIGUCHI M.;  
RT "HLA-B51 and HLA-Bw52 differ by only two amino acids which are in the  
helical region of the alpha 1 domain.";  
RL J. Immunol. 142:306-311(1989).  
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO  
THE IMMUNE SYSTEM.  
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
MICROGLOBULIN).  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC ENBL; M22799; AAA59645.1; ALT\_SEQ.  
DR ENBL; M22793; AAA59645.1; JOINED.  
DR ENBL; M22794; AAA59645.1; JOINED.  
DR ENBL; M22795; AAA59645.1; JOINED.  
DR ENBL; M22796; AAA59645.1; JOINED.  
DR ENBL; M22797; AAA59645.1; JOINED.  
DR ENBL; M22798; AAA59645.1; JOINED.  
DR PIR; B30345; B30345.  
DR PIR; B30348; B30348.  
DR HSSP; P30491; IALM.  
DR MIM; 142830;  
DR PROSITE; PS00290; IG\_MHC; 1.  
DR PFAM; PF00047; ig; 1.  
DR PFAM; PF00129; MHC-I; 1.  
KW MHC I; Transmembrane; Glycoprotein; Signal.  
FT SIGNAL 1 24

FT CHAIN 25 362 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,  
BW-52(B-5) B\*5201 ALPHA CHAIN.  
FT DOMAIN 25 114 EXTRACELLULAR ALPHA-1.  
FT DOMAIN 115 206 EXTRACELLULAR ALPHA-2.  
FT DOMAIN 207 298 EXTRACELLULAR ALPHA-3.  
FT DOMAIN 299 308 CONNECTING PEPTIDE.  
FT TRANSMEM 309 332  
FT DOMAIN 333 362 CYTOPLASMIC TAIL.  
FT CARBOHYD 110 110 BY SIMILARITY.  
FT DISULFID 125 188 BY SIMILARITY.  
FT DISULFID 227 283 BY SIMILARITY.  
SQ SEQUENCE 362 AA; 40521 MW; 3B436FE8 CRC32;

Query Match 50.5%; Score 50; DB 1; Length 362;  
Best Local Similarity 68.8%; Pred. No. 0.46;  
Matches 11; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 5 TRLNERENLRALRY 20  
I | | | | | | | | | |  
Db 93 TMTQTYRENLRALRY 108

RESULT 5  
1B54\_HUMAN  
ID 1B54\_HUMAN STANDARD; PRT; 362 AA.  
AC P30491;  
DT 01-APR-1993 (Rel. 25, Created)  
DT 01-APR-1993 (Rel. 25, Last sequence update)  
DT 13-JUL-1998 (Rel. 36, Last annotation update)  
DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, BW-53 B\*5301 ALPHA CHAIN  
DE PRECURSOR.  
GN HLA-B OR HLAB.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Homnidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 91033941.  
RA HAYASHI H., Ooba T., NAKAYAMA S., SEKIMATA M., KANO K.,  
TAKIGUCHI M.;  
RT "Allspecificities between HLA-Bw53 and HLA-B35 are generated by  
substitution of the residues associated with HLA-Bw4/Bw6 public  
epitopes.";  
RL Immunogenetics 32:195-199(1990).  
RN [2]  
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 25-302.  
RX MEDLINE; 96209672.  
RA SMITH K.J., REID S.W., HARLOS K., MCMICHAEL A.J., STUART D.I.,  
BELL J.I., JONES E.Y.;  
RT "Bound water structure and polymorphic amino acids act together to  
allow the binding of different peptides to MHC class I HLA-B53.";  
RL Immunity 4:215-228(1996).  
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO  
THE IMMUNE SYSTEM.  
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
MICROGLOBULIN).  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC ENBL; M58636; AAA36228.1;  
DR PIR; A45834; A45834.  
DR PDB; 1ALM; 08-APR-98.  
DR PDB; 1AIO; 08-APR-98.  
DR MIM; 142830;  
DR PROSITE; PS00290; IG\_MHC; 1.  
DR PFAM; PF00047; ig; 1.

DR PFAM; PF00129; MHC\_I; 1.  
 KW MHC I; Transmembrane; Glycoprotein; Signal; 3D-structure.  
 FT SIGNAL 1 24  
 FT CHAIN 25 362 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,  
 FT BW-53 B\*5301 ALPHA CHAIN.  
 FT DOMAIN 25 114 EXTRACELLULAR ALPHA-1.  
 FT DOMAIN 115 206 EXTRACELLULAR ALPHA-2.  
 FT DOMAIN 207 298 EXTRACELLULAR ALPHA-3.  
 FT DOMAIN 299 308 CONNECTING PEPTIDE.  
 FT TRANSMEM 309 332 CYTOPLASMIC TAIL.  
 FT CARBOHYD 333 362 BY SIMILARITY.  
 FT DISULFID 110 110  
 FT DISULFID 125 188  
 FT SEQUENCE 362 AA; 40495 MW; 2BDC746E CRC32;  
 SQ

Query Match 50.5%; Score 50; DB 1; Length 362;  
 Best Local Similarity 68.8%; Pred. No. 0.46;  
 Matches 11; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 TRLNERENLRALRY 20  
 DB 93 TNRQTYRENLRALRY 108

RESULT 6  
 ID 1B01\_PANTR STANDARD; PRT: 359 AA.  
 AC P13750;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 01-APR-1993 (Rel. 25, Last annotation update)  
 DE CHLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-1 ALPHA CHAIN PRECURSOR  
 DE (FRAGMENT).  
 OS Pan troglodytes (Chimpanzee).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Pan.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 89030641.  
 RA MAYER W.E., JONKER M., KLEIN D., IVANYI P., VAN SEVENTER G.,  
 RA KLEIN J.;  
 RT "Nucleotide sequences of chimpanzee MHC class I alleles: evidence for  
 RT trans-species mode of evolution.";  
 RL EMBO J. 7:2765-2774(1988).  
 RN [2]  
 RP REVISIONS.  
 RA MAYER W.;  
 RL Submitted (FEB-1989) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO  
 CC THE IMMUNE SYSTEM.  
 CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
 CC MICROGLOBULIN).  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL; X13115; CAA31507.1; -  
 DR PIR; S03537; S03537.  
 DR HSSP; P03989; ILSA.  
 DR PROSITE; PS00290; IG\_MHC; 1.  
 DR PFAM; PF00047; ig; 1.  
 DR MHC I; Transmembrane; Glycoprotein; Signal.  
 KW MHC I; Transmembrane; Glycoprotein; Signal.  
 FT NON\_TER 1  
 FT SIGNAL <1 20  
 FT CHAIN 21 359 CHLA CLASS I HISTOCOMPATIBILITY ANTIGEN,  
 FT

FT DOMAIN 21 110 B-1 ALPHA CHAIN.  
 FT DOMAIN 111 202 EXTRACELLULAR ALPHA-1.  
 FT DOMAIN 203 294 EXTRACELLULAR ALPHA-2.  
 FT DOMAIN 295 305 EXTRACELLULAR ALPHA-3.  
 FT TRANSMEM 306 329 CONNECTING PEPTIDE.  
 FT DOMAIN 330 359 CYTOPLASMIC TAIL.  
 FT DISULFID 121 184 BY SIMILARITY.  
 FT DISULFID 223 279 BY SIMILARITY.  
 FT CARBOHYD 106 106  
 FT SEQUENCE 359 AA; 40173 MW; 5395FFC9 CRC32;  
 SQ

Query Match 49.5%; Score 49; DB 1; Length 359;  
 Best Local Similarity 100.0%; Pred. No. 0.67;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RENLRALRY 20  
 DB 95 RENLRALRY 104

RESULT 7  
 ID 1B01\_GORGO STANDARD; PRT: 362 AA.  
 AC P30379;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 01-APR-1993 (Rel. 25, Last annotation update)  
 DE CLASS I HISTOCOMPATIBILITY ANTIGEN, GOGO-B0101 ALPHA CHAIN PRECURSOR.  
 OS Gorilla gorilla gorilla (Lowland gorilla).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Gorilla.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 92078860.  
 RA LAWLOR D.A., WARREN E., TAYLOR P., PARHAM P.;  
 RT "Gorilla class I major histocompatibility complex alleles: comparison  
 RT to human and chimpanzee class I";  
 RL J. Exp. Med. 174:1491-1509(1991).  
 CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO  
 CC THE IMMUNE SYSTEM.  
 CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
 CC MICROGLOBULIN).  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL; X60255; CAA42807.1; -  
 DR PIR; JH0539; JH0539.  
 DR HSSP; P03989; ILSA.  
 DR PROSITE; PS00290; IG\_MHC; 1.  
 DR PFAM; PF00047; ig; 1.  
 DR MHC I; Transmembrane; Glycoprotein; Signal.  
 KW MHC I; Transmembrane; Glycoprotein; Signal.  
 FT SIGNAL 1 24  
 FT CHAIN 25 362 CLASS I HISTOCOMPATIBILITY ANTIGEN,  
 FT GOGO-B0101 ALPHA CHAIN.  
 FT DOMAIN 25 114 EXTRACELLULAR ALPHA-1.  
 FT DOMAIN 115 206 EXTRACELLULAR ALPHA-2.  
 FT DOMAIN 207 298 EXTRACELLULAR ALPHA-3.  
 FT DOMAIN 299 308 CONNECTING PEPTIDE.  
 FT TRANSMEM 309 332 CYTOPLASMIC TAIL.  
 FT DOMAIN 333 362 BY SIMILARITY.  
 FT DISULFID 125 188 BY SIMILARITY.  
 FT DISULFID 227 283 BY SIMILARITY.  
 FT CARBOHYD 110 110  
 FT SEQUENCE 362 AA; 40170 MW; 2E33E2B8 CRC32;  
 SQ

Query Match 49.5%; Score 49; DB 1; Length 362;  
 Best Local Similarity 100.0%; Pred. No. 0.67;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RENLIRALRY 20  
 |||||  
 Db 99 RENLIRALRY 108

RESULT 8  
 1B02\_GORGO STANDARD; PRT; 362 AA.  
 AC P30380;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DE CLASS I HISTOCOMPATIBILITY ANTIGEN, GOGO-B0102 ALPHA CHAIN PRECURSOR.  
 OS Gorilla gorilla gorilla (Lowland gorilla).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Gorilla.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 92078860.  
 RA LAWLOR D.A., WARREN E., TAYLOR P., PARHAM P.;  
 RT "Gorilla class I major histocompatibility complex alleles: comparison  
 to human and chimpanzee class I.";  
 RL J. Exp. Med. 174:1491-1509(1991).  
 CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO  
 THE IMMUNE SYSTEM.  
 CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
 MICROGLOBULIN).  
 CC  
 CC SEQUENCE FROM N.A.  
 RX MEDLINE; 92078860.  
 RA LAWLOR D.A., WARREN E., TAYLOR P., PARHAM P.;  
 RT "Gorilla class I major histocompatibility complex alleles: comparison  
 to human and chimpanzee class I.";  
 RL J. Exp. Med. 174:1491-1509(1991).  
 CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO  
 THE IMMUNE SYSTEM.  
 CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
 MICROGLOBULIN).  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC

EMBL; X60693; CAA43101.1; -  
 DR PIR; JH0540; JH0540.  
 DR HSSP; P03989; ILSA.  
 DR PROSITE; PS00290; IG\_MHC; 1.  
 DR PFAM; PF00047; ig; 1.  
 DR MHC I; Transmembrane; Glycoprotein; Signal.  
 FT SIGNAL 1 24  
 FT CHAIN 25 362  
 FT DOMAIN 25 114  
 FT DOMAIN 115 206  
 FT DOMAIN 207 298  
 FT DOMAIN 299 308  
 FT TRANSMEM 309 332  
 FT TRANSMEM 333 362  
 FT DOMAIN 333 362  
 FT DISULFID 125 188  
 FT DISULFID 227 283  
 FT CARBOHYD 110 110  
 SQ SEQUENCE 362 AA; 40204 MW; 3CF119AD CRC32;

Query Match 49.5%; Score 49; DB 1; Length 362;  
 Best Local Similarity 100.0%; Pred. No. 0.67;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RENLIRALRY 20  
 |||||  
 Db 99 RENLIRALRY 108

RESULT 9  
 1B03\_GORGO STANDARD; PRT; 362 AA.  
 AC P30381;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DE CLASS I HISTOCOMPATIBILITY ANTIGEN, GOGO-B0103 ALPHA CHAIN PRECURSOR.  
 OS Gorilla gorilla gorilla (Lowland gorilla).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Gorilla.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 92078860.  
 RA LAWLOR D.A., WARREN E., TAYLOR P., PARHAM P.;  
 RT "Gorilla class I major histocompatibility complex alleles: comparison  
 to human and chimpanzee class I.";  
 RL J. Exp. Med. 174:1491-1509(1991).  
 CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO  
 THE IMMUNE SYSTEM.  
 CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
 MICROGLOBULIN).  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC

EMBL; X60254; CAA42806.1; -  
 DR PIR; JH0541; JH0541.  
 DR HSSP; P03989; ILSA.  
 DR PROSITE; PS00290; IG\_MHC; 1.  
 DR PFAM; PF00047; ig; 1.  
 DR MHC I; Transmembrane; Glycoprotein; Signal.  
 FT SIGNAL 1 24  
 FT CHAIN 25 362  
 FT DOMAIN 25 114  
 FT DOMAIN 115 206  
 FT DOMAIN 207 298  
 FT DOMAIN 299 308  
 FT TRANSMEM 309 332  
 FT TRANSMEM 333 362  
 FT DOMAIN 333 362  
 FT DISULFID 125 188  
 FT DISULFID 227 283  
 FT CARBOHYD 110 110  
 SQ SEQUENCE 362 AA; 40248 MW; FEA6A941 CRC32;

Query Match 49.5%; Score 49; DB 1; Length 362;  
 Best Local Similarity 100.0%; Pred. No. 0.67;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RENLIRALRY 20  
 |||||  
 Db 99 RENLIRALRY 108

RESULT 10  
 1B15\_HUMAN STANDARD; PRT; 362 AA.  
 AC P10317;  
 DT 01-MAR-1989 (Rel. 10, Created)  
 DT 01-MAR-1989 (Rel. 10, Last sequence update)  
 DE CLASS I HISTOCOMPATIBILITY ANTIGEN, B-27 B\*2702 ALPHA CHAIN  
 PRECURSOR (B-27K) (B27.2).  
 GN HLA-B OR HLAB.  
 OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RN SEQUENCE FROM N.A.  
RX MEDLINE: 86220133.  
RA SEEMANN G.H.A., REIN R.S., BROWN C.S., PLOEGH H.L.;  
RT "Gene conversion-like mechanisms may generate polymorphism in human  
RT class I genes";  
RL EMBO J. 5:547-552(1986).  
RN [2]  
RN SEQUENCE FROM N.A.  
RP PARHAM P., ARNETT K.L., ADAMS E.J.;  
RA Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RX SEQUENCE OF 86-107 AND 171-181.  
RA VEGA M.A., EZQUERRA A., ROJO S., APARTICIO P., BRAGADO R.,  
RA LOPEZ DE CASTRO J.A.;  
RT "Structural analysis of an HLA-B27 functional variant: identification  
RT of residues that contribute to the specificity of recognition by  
RT cytolytic T lymphocytes";  
RL Proc. Natl. Acad. Sci. U.S.A. 82:7394-7398(1985).  
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO  
CC THE IMMUNE SYSTEM.  
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
CC MICROGLOBULIN).  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
DR EMBL: X03664; CAA27301.1; -;  
DR EMBL: X03667; CAA27301.1; JOINED.  
DR EMBL: L38504; AAG69724.1; -;  
DR PIR: B25092; HLHUBK.  
DR HSP: P03989; LHSN.  
DR MIM: 142830; -;  
DR PROSITE: PS00290; IG\_MHC; 1.  
DR PFAM: PF00047; ig; 1.  
DR PFAM: PF00129; MHC\_I; 1.  
KW MHC I; Transmembrane; Glycoprotein; Signal.  
FT SIGNAL 1 24  
FT CHAIN 25 362 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,  
FT B-27 B\*2702 ALPHA CHAIN.  
FT DOMAIN 25 114  
FT DOMAIN 115 206 EXTRACELLULAR ALPHA-1.  
FT DOMAIN 207 298 EXTRACELLULAR ALPHA-2.  
FT DOMAIN 299 308 EXTRACELLULAR ALPHA-3.  
FT DOMAIN 309 332 CONNECTING PEPTIDE.  
FT TRANSMEM 309 332  
FT DOMAIN 333 362 CYTOPLASMIC TAIL.  
FT CARBOHYD 110 110 BY SIMILARITY.  
FT DISULFID 125 188 BY SIMILARITY.  
FT DISULFID 227 283 BY SIMILARITY.  
SQ SEQUENCE 362 AA; 40397 MW; 9798F0BB CRC32;

Query Match 49.5%; Score 49; DB 1; Length 362;  
Best Local Similarity 100.0%; Pred. No. 0.67;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RENLRLALRY 20  
DB 99 RENLRLALRY 108  
|||||

RESULT 11  
1B60\_HUMAN STANDARD; PRT; 362 AA.  
ID 1B60\_HUMAN  
AC P18465;

DT 01-NOV-1990 (Rel. 16, Created)  
DT 01-NOV-1990 (Rel. 16, Last sequence update)  
DT 01-APR-1993 (Rel. 25, Last annotation update)  
DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-57(B-17) B\*5701 ALPHA  
DE CHAIN PRECURSOR (BWS7.1).  
GN HLA-B OR HLAB.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RN SEQUENCE FROM N.A.  
RX MEDLINE: 90207291.  
RA ENNIS P.D., ZEMMOUR J., SALTER R.D., PARHAM P.;  
RT "Rapid cloning of HLA-A,B CDNA by using the polymerase chain  
RT reaction: frequency and nature of errors produced in amplification";  
RL Proc. Natl. Acad. Sci. U.S.A. 87:2833-2837(1990).  
RN [2]  
RN SEQUENCE FROM N.A.  
RX MEDLINE: 91067476.  
RA ISAMAT M., GIRDLESTONE J., MILSTEIN C.;  
RT Nucleotide sequence of an HLA-Bw57 gene";  
RL Nucleic Acids Res. 18:6702-6702(1990).  
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO  
CC THE IMMUNE SYSTEM.  
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
CC MICROGLOBULIN).  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
DR EMBL: M32318; AAA36231.1; -;  
DR EMBL: X55711; CAA39244.1; -;  
DR PIR: S12622; S12622.  
DR PIR: D35997; D35997.  
DR HSP: P30491; IAIM.  
DR MIM: 142830; -;  
DR PROSITE: PS00290; IG\_MHC; 1.  
DR PFAM: PF00047; ig; 1.  
DR PFAM: PF00129; MHC\_I; 1.  
KW MHC I; Transmembrane; Glycoprotein; Signal.  
FT SIGNAL 1 24  
FT CHAIN 25 362 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,  
FT B-57(B-17) B\*5701 ALPHA CHAIN.  
FT DOMAIN 25 114  
FT DOMAIN 115 206 EXTRACELLULAR ALPHA-1.  
FT DOMAIN 207 298 EXTRACELLULAR ALPHA-2.  
FT DOMAIN 299 308 EXTRACELLULAR ALPHA-3.  
FT DOMAIN 309 332 CONNECTING PEPTIDE.  
FT TRANSMEM 309 332  
FT DOMAIN 333 362 CYTOPLASMIC TAIL.  
FT CARBOHYD 110 110 BY SIMILARITY.  
FT DISULFID 125 188 BY SIMILARITY.  
FT DISULFID 227 283 BY SIMILARITY.  
SQ SEQUENCE 362 AA; 40224 MW; D91DF8DD CRC32;

Query Match 49.5%; Score 49; DB 1; Length 362;  
Best Local Similarity 100.0%; Pred. No. 0.67;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RENLRLALRY 20  
DB 99 RENLRLALRY 108  
|||||

RESULT 12  
1B61\_HUMAN STANDARD; PRT; 362 AA.  
ID 1B61\_HUMAN  
AC P30497;

DT 01-APR-1993 (Rel. 25, Created)  
DT 01-APR-1993 (Rel. 25, Last sequence update)  
DT 01-APR-1993 (Rel. 25, Last annotation update)  
DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B\*57(B-17) B\*5702 ALPHA CHAIN  
DE PRECURSOR (BW57.2).  
GN HLA-B OR HLAB.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Homiidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 93056508.  
RA MADRIGAL J.A., ZEMMOUR J., BELICH M.P., HILDEBRAND W.H., BENJAMIN R.J.,  
RA LITTLE A.-M., ZEMMOUR J., ENNIS P.D., WARD F.E., PETZL-ERLER M.L.,  
RA MARTIN R.W., DU TOIT E.D., PARHAM P.;  
RT "Distinctive HLA-A,B antigens of black populations formed by  
RT interallelic conversion";  
RL J. Immunol. 149:3411-3415(1992).  
CC -1- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO  
CC THE IMMUNE SYSTEM.  
CC -1- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
CC MICROGLOBULIN).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X61707; CAA43876.1; -  
DR PIR; S16774; S16774.  
DR HSSP; P30491; 1ALM.  
DR MIM; 142830; -  
DR PROSITE; PS00290; IG\_MHC; 1.  
DR PFAM; PF00047; IG; 1.  
DR PFAM; PF00129; MHC\_I; 1.  
KW MHC I; Transmembrane; Glycoprotein; Signal.  
FT SIGNAL 1 24  
FT CHAIN 25 362 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,  
FT DOMAIN 25 362 B\*57(B-17) B\*5702 ALPHA CHAIN.  
FT DOMAIN 25 114 EXTRACELLULAR ALPHA-1.  
FT DOMAIN 115 206 EXTRACELLULAR ALPHA-2.  
FT DOMAIN 207 298 EXTRACELLULAR ALPHA-3.  
FT DOMAIN 299 308 CONNECTING PEPTIDE.  
FT TRANSMEM 309 332  
FT DOMAIN 333 362 CYTOPLASMIC TAIL.  
FT CARBOHYD 110 110 BY SIMILARITY.  
FT DISULFID 125 188 BY SIMILARITY.  
FT DISULFID 227 283 BY SIMILARITY.  
SQ SEQUENCE 362 AA; 40342 MW; 628C2156 CRC32;

Query Match 49.5%; Score 49; DB 1; Length 362;  
Best Local Similarity 100.0%; Pred. No. 0.67;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 11 RENLIRALRY 20  
DB 99 RENLIRALRY 108

RESULT 13  
1B62\_HUMAN  
ID 1B62\_HUMAN STANDARD; PRT; 362 AA.  
AC P10319;  
DT 01-MAR-1989 (Rel. 10, Created)  
DT 01-MAR-1989 (Rel. 10, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, BW-56(B-17) B\*5801 ALPHA  
DE CHAIN PRECURSOR.  
GN HLA-B OR HLAB.

OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Homiidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 86008247.  
RA WAYS J.P., COPPIN H.L., PARHAM P.;  
RT "The complete primary structure of HLA-B\*58.";  
RL J. Biol. Chem. 260:11924-11933(1985).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 86008247.  
RA INOUE T., OGAWA A.;  
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO  
CC THE IMMUNE SYSTEM.  
CC -1- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
CC MICROGLOBULIN).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M11799; AAA59628.1; -  
DR EMBL; AB008102; BAA22916.1; -  
DR PIR; A23895; HLHUB8.  
DR HSSP; P30491; 1ALM.  
DR MIM; 142830; -  
DR PROSITE; PS00290; IG\_MHC; 1.  
DR PFAM; PF00047; IG; 1.  
DR PFAM; PF00129; MHC\_I; 1.  
KW MHC I; Transmembrane; Glycoprotein; Signal.  
FT SIGNAL 1 24  
FT CHAIN 25 362 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,  
FT DOMAIN 25 114 EXTRACELLULAR ALPHA-1.  
FT DOMAIN 115 206 EXTRACELLULAR ALPHA-2.  
FT DOMAIN 207 298 EXTRACELLULAR ALPHA-3.  
FT DOMAIN 299 308 CONNECTING PEPTIDE.  
FT TRANSMEM 309 332  
FT DOMAIN 333 362 CYTOPLASMIC TAIL.  
FT CARBOHYD 110 110 BY SIMILARITY.  
FT DISULFID 125 188 BY SIMILARITY.  
FT DISULFID 227 283 BY SIMILARITY.  
SQ SEQUENCE 362 AA; 40337 MW; 3E5E7534 CRC32;

Query Match 49.5%; Score 49; DB 1; Length 362;  
Best Local Similarity 100.0%; Pred. No. 0.67;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 11 RENLIRALRY 20  
DB 99 RENLIRALRY 108

RESULT 14  
HLAH\_HUMAN  
ID HLAH\_HUMAN STANDARD; PRT; 362 AA.  
AC P01893;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 01-APR-1993 (Rel. 25, Last sequence update)  
DT 15-DEC-1999 (Rel. 39, Last annotation update)  
DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, ALPHA CHAIN H PRECURSOR  
DE (HLA-AR) (HLA-12.4).  
GN HLA-H OR HLAH.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Homiidae; Homo.

```

RN  SEQUENCE FROM N.A.
RX  MEDLINE; 82151002.
RA  MALISSEN M., MALISSEN B., JORDAN B.R.;
RT  "Exon/intron organization and complete nucleotide sequence of an HLA
RL  gene.";
CC  Proc. Natl. Acad. Sci. U.S.A. 79:893-897(1982).
CC  -1- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC  THE IMMUNE SYSTEM. COULD BE THE PRODUCT OF A PSEUDOGENE.
CC  -1- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC  MICROGLOBULIN).
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
CC  EMBL; J00191; AAA36218.1; ALT_INIT.
CC  PIR; A02189; HLH12.
CC  HSP; P03989; IHSA.
CC  MIN; 142800.
CC  PROSITE; PS00290; IG_MHC; 1.
CC  PFAM; PF00047; Ig; 1.
CC  PFAM; PF00129; MHC_I; 1.
CC  MHC_I; Transmembrane; Glycoprotein; Signal.
CC  SIGNAL 1 24
CC  CHAIN 25 362
CC  FT FT HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
CC  FT FT ALPHA CHAIN H.
CC  FT FT EXTRACELLULAR ALPHA-1.
CC  FT FT EXTRACELLULAR ALPHA-2.
CC  FT FT EXTRACELLULAR ALPHA-3.
CC  FT FT CONNECTING PEPTIDE.
CC  FT FT TRANSMEM
CC  FT FT CYTOPLASMIC TAIL.
CC  FT FT BY SIMILARITY.
CC  FT FT DISULFID 227 283
CC  FT FT BY SIMILARITY.
CC  FT FT SEQUENCE 362 AA; 40850 MW; 5E610F63 CRC32;
CC  -----
CC  Query Match 49.5%; Score 49; DB 1; Length 362;
CC  Best Local Similarity 100.0%; Pred. No. 0.67;
CC  Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC  Qy 11 RENLIRALRY 20
CC  Db 99 RENLIRALRY 108
CC  -----
RESULT 15
1A23_HUMAN
ID 1A23_HUMAN STANDARD; PRT; 365 AA.
AC P30447;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, A-23(A-9) ALPHA CHAIN
DE PRECURSOR.
DE HLA-A OR HLA-A.
GN Homo sapiens (Human).
OS Chordata; Craniata; Vertebrata; Mammalia;
OC Eukaryota; Metazoa;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
CC [1]
CC SEQUENCE FROM N.A. (A*2301).
CC MEDLINE; 92104637.
CC LITTLE A.-M., MADRIGAL J.A., PARHAM P.;
CC "Molecular definition of an elusive third HLA-A9 molecule: HLA-A9.3.";
CC Immunogenetics 35:41-45(1992).
CC -1- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC THE IMMUNE SYSTEM.
CC -1- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
```

```

CC MICROGLOBULIN).
CC -1- POLYMORPHISM: THE ONLY ALLELE OF A-23 KNOWN IS A*2301 WHICH IS
CC SHOWN HERE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M64742; AAA03662.1; -.
CC HSP; P01892; IAQD.
CC MIN; 142800.
CC PROSITE; PS00290; IG_MHC; 1.
CC PFAM; PF00047; Ig; 1.
CC PFAM; PF00129; MHC_I; 1.
CC MHC_I; Transmembrane; Glycoprotein; Signal.
CC SIGNAL 1 24
CC CHAIN 25 365
CC FT FT HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
CC FT FT A-23(A-9) ALPHA CHAIN.
CC FT FT EXTRACELLULAR ALPHA-1.
CC FT FT EXTRACELLULAR ALPHA-2.
CC FT FT EXTRACELLULAR ALPHA-3.
CC FT FT CONNECTING PEPTIDE.
CC FT FT TRANSMEM
CC FT FT CYTOPLASMIC TAIL.
CC FT FT BY SIMILARITY.
CC FT FT DISULFID 110 110
CC FT FT BY SIMILARITY.
CC FT FT SEQUENCE 365 AA; 40732 MW; B1C21094 CRC32;
CC -----
CC Query Match 49.5%; Score 49; DB 1; Length 365;
CC Best Local Similarity 100.0%; Pred. No. 0.68;
CC Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC Qy 11 RENLIRALRY 20
CC Db 99 RENLIRALRY 108
CC -----
Search completed: February 8, 2000, 01:25:59
Job time: 1559 sec
```



Result No.	Score	Query %		Length	DB	ID	Description
		Match					
1	50	50.5	89	7	019569	019569	homo sapien
2	50	50.5	172	7	019770	019770	homo sapien
3	50	50.5	172	7	019774	019774	homo sapien
4	50	50.5	172	7	019775	019775	homo sapien
5	50	50.5	172	7	019780	019780	homo sapien
6	50	50.5	172	7	095364	095364	homo sapien
7	50	50.5	172	7	019771	019771	homo sapien
8	50	50.5	172	7	019772	019772	homo sapien
9	50	50.5	172	7	019773	019773	homo sapien
10	50	50.5	175	7	029694	029694	homo sapien
11	50	50.5	181	7	046703	046703	homo sapien
12	50	50.5	181	7	062899	062899	homo sapien
13	50	50.5	181	7	062922	062922	homo sapien
14	50	50.5	181	7	019623	019623	homo sapien
15	50	50.5	181	7	019747	019747	homo sapien
16	50	50.5	181	7	029667	029667	homo sapien
17	50	50.5	181	7	030198	030198	homo sapien
18	50	50.5	181	7	029708	029708	homo sapien
19	50	50.5	181	7	P79559	P79559	homo sapien
20	50	50.5	181	7	019597	019597	homo sapien

GN HLA-B.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

OC Eutheria; Primates; Catarrhini; Hominidae; Homo.

RN [1]  
RP SEQUENCE FROM N.A.  
RA CHANDANAYINGYONG D., SIRIKONG M., LONGTA K., SRINAK D., RUNGROUNG E.,  
RA BEJCHANDRA S., JUJI T., TOKUNAGA K.;  
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U90421; AAB50144.1; -.  
DR EMBL; U90420; AAB50144.1; JOINED.  
DR PFAM; PF00129; MHC\_I; 1.  
KW MHC.  
FT NON\_TER 1  
FT NON\_TER 172 172  
SQ SEQUENCE 172 AA; 19909 MW; CAAE5641 CRC32;

Query Match 50.5%; Score 50; DB 7; Length 172;  
Best Local Similarity 68.8%; Pred. No. 1.3;  
Matches 11; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 TRLNERENLRALRY 20  
DB 62 TNTQTYENLRALRY 77

RESULT 3

ID O19774 PRELIMINARY; PRT; 172 AA.  
AC O19774;  
DT 01-JAN-1998 (TEMBLrel. 05, Created)  
DT 01-JAN-1998 (TEMBLrel. 05, Last sequence update)  
DT 01-NOV-1998 (TEMBLrel. 08, Last annotation update)  
DE MHC CLASS I HLA-B (FRAGMENT).  
GN HLA-B.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA CHANDANAYINGYONG D., SIRIKONG M., LONGTA K., SRINAK D., RUNGROUNG E.,  
RA BEJCHANDRA S., BLASZYK R., GROSE-WILDE H.;  
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U90423; AAB50145.1; -.  
DR EMBL; U90422; AAB50145.1; JOINED.  
DR PFAM; PF00129; MHC\_I; 1.  
KW MHC.  
FT NON\_TER 1  
FT NON\_TER 172 172  
SQ SEQUENCE 172 AA; 19909 MW; CAAE5641 CRC32;

Query Match 50.5%; Score 50; DB 7; Length 172;  
Best Local Similarity 68.8%; Pred. No. 1.3;  
Matches 11; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 TRLNERENLRALRY 20  
DB 62 TNTQTYENLRALRY 77

RESULT 4

ID O19775 PRELIMINARY; PRT; 172 AA.  
AC O19775;  
DT 01-JAN-1998 (TEMBLrel. 05, Created)  
DT 01-JAN-1998 (TEMBLrel. 05, Last sequence update)  
DT 01-NOV-1998 (TEMBLrel. 08, Last annotation update)  
DE MHC CLASS I HLA-B (FRAGMENT).  
GN HLA-B.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.

RA CHANDANAYINGYONG D., SIRIKONG M., LONGTA K., SRINAK D., RUNGROUNG E.,  
RA BEJCHANDRA S.;  
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U90425; AAB50146.1; -.  
DR EMBL; U90424; AAB50146.1; JOINED.  
DR PFAM; PF00129; MHC\_I; 1.  
KW MHC.  
FT NON\_TER 1  
FT NON\_TER 172 172  
SQ SEQUENCE 172 AA; 19909 MW; CAAE5641 CRC32;

Query Match 50.5%; Score 50; DB 7; Length 172;  
Best Local Similarity 68.8%; Pred. No. 1.3;  
Matches 11; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 TRLNERENLRALRY 20  
DB 62 TNTQTYENLRALRY 77

RESULT 5

ID O19780 PRELIMINARY; PRT; 172 AA.  
AC O19780;  
DT 01-JAN-1998 (TEMBLrel. 05, Created)  
DT 01-JAN-1998 (TEMBLrel. 05, Last sequence update)  
DT 01-NOV-1998 (TEMBLrel. 08, Last annotation update)  
DE MHC CLASS I HLA-B (FRAGMENT).  
GN HLA-B.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA CHANDANAYINGYONG D., SIRIKONG M., LONGTA K., SRINAK D., RUNGROUNG E.,  
RA BEJCHANDRA S., JUJI T., TOKUNAGA K.;  
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U90419; AAB50143.1; -.  
DR EMBL; U90418; AAB50143.1; JOINED.  
DR PFAM; PF00129; MHC\_I; 1.  
KW MHC.  
FT NON\_TER 1  
FT NON\_TER 172 172  
SQ SEQUENCE 172 AA; 19909 MW; CAAE5641 CRC32;

Query Match 50.5%; Score 50; DB 7; Length 172;  
Best Local Similarity 68.8%; Pred. No. 1.3;  
Matches 11; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 TRLNERENLRALRY 20  
DB 62 TNTQTYENLRALRY 77

RESULT 6

ID O95364 PRELIMINARY; PRT; 172 AA.  
AC O95364;  
DT 01-FEB-1997 (TEMBLrel. 02, Created)  
DT 01-FEB-1997 (TEMBLrel. 02, Last sequence update)  
DT 01-NOV-1998 (TEMBLrel. 08, Last annotation update)  
DE MHC HLA-B\*51 PROTEIN (FRAGMENT).  
GN HLA-B\*51FA.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA BLASZYK R.;  
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.  
DR EMBL; X96473; CAA65327.1; -.

DR PFAM: PF00129; MHC\_I; 1.  
KW MHC.

FT NON\_TER 172 172

SQ SEQUENCE 172 AA; 19942 MW; 1A73E47D CRC32;

Query Match 50.5%; Score 50; DB 7; Length 172;  
Best Local Similarity 68.8%; Pred. No. 1.3;  
Matches 11; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 TRLNERRENLRALRY 20

Db 59 TNTQTYRENLRALRY 74

RESULT 7

019771

ID 019771 PRELIMINARY; PRT; 172 AA.

AC 019771;

DT 01-JAN-1998 (TREMELrel. 05, Created)

DT 01-JAN-1998 (TREMELrel. 05, Last sequence update)

DT 01-NOV-1998 (TREMELrel. 08, Last annotation update)

DE MHC CLASS I HLA-B (FRAGMENT).

GN HLA-B.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

OC Eutheria; Primates; Catarrhini; Hominidae; Homo.

RN [1]

RA CHANDANAYONG D., SIRIKONG M., LONGTA K., SRINAK D.,

RA SIRIBOONRIT U., RUNGROUNG E., BEJCHANDRA S.;

RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.

DR EMBL: U90612; AAB50151.1;

DR EMBL: U90611; AAB50151.1; JOINED.

DR PFAM: PF00129; MHC\_I; 1.

KW MHC.

FT NON\_TER 172 172

SQ SEQUENCE 172 AA; 20026 MW; 4D9A1043 CRC32;

Query Match

Best Local Similarity 50.5%; Score 50; DB 7; Length 172;  
Matches 11; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 TRLNERRENLRALRY 20

Db 62 TNTQTYRENLRALRY 77

RESULT 8

019772

ID 019772 PRELIMINARY; PRT; 172 AA.

AC 019772;

DT 01-JAN-1998 (TREMELrel. 05, Created)

DT 01-JAN-1998 (TREMELrel. 05, Last sequence update)

DT 01-NOV-1998 (TREMELrel. 08, Last annotation update)

DE MHC CLASS I HLA-B (FRAGMENT).

GN HLA-B.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

OC Eutheria; Primates; Catarrhini; Hominidae; Homo.

RN [1]

RA CHANDANAYONG D., SIRIKONG M., LONGTA K., SRINAK D.,

RA SIRIBOONRIT U., RUNGROUNG E., BEJCHANDRA S.;

RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.

DR EMBL: U90614; AAB50244.1;

DR EMBL: U90613; AAB50244.1; JOINED.

DR PFAM: PF00129; MHC\_I; 1.

KW MHC.

FT NON\_TER 172 172

FT NON\_TER 172 172

SQ SEQUENCE 172 AA; 20026 MW; 4D9A1043 CRC32;

Query Match 50.5%; Score 50; DB 7; Length 172;  
Best Local Similarity 68.8%; Pred. No. 1.3;  
Matches 11; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 TRLNERRENLRALRY 20

Db 62 TNTQTYRENLRALRY 77

RESULT 9

019773

ID 019773 PRELIMINARY; PRT; 172 AA.

AC 019773;

DT 01-JAN-1998 (TREMELrel. 05, Created)

DT 01-JAN-1998 (TREMELrel. 05, Last sequence update)

DT 01-NOV-1998 (TREMELrel. 08, Last annotation update)

DE MHC CLASS I HLA-B (FRAGMENT).

GN HLA-B.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

OC Eutheria; Primates; Catarrhini; Hominidae; Homo.

RN [1]

RA CHANDANAYONG D., SIRIKONG M., LONGTA K., SRINAK D.,

RA SIRIBOONRIT U., RUNGROUNG E., BEJCHANDRA S.;

RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.

DR EMBL: U90616; AAB50245.1;

DR EMBL: U90615; AAB50245.1; JOINED.

DR PFAM: PF00129; MHC\_I; 1.

KW MHC.

FT NON\_TER 172 172

SQ SEQUENCE 172 AA; 20052 MW; F6214671 CRC32;

Query Match

Best Local Similarity 50.5%; Score 50; DB 7; Length 172;  
Matches 11; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 TRLNERRENLRALRY 20

Db 62 TNTQTYRENLRALRY 77

RESULT 10

Q29694

ID Q29694 PRELIMINARY; PRT; 175 AA.

AC Q29694;

DT 01-NOV-1996 (TREMELrel. 01, Created)

DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)

DT 01-NOV-1996 (TREMELrel. 12, Last annotation update)

DE MHC CLASS I HLA-B ANTIGEN (FRAGMENT).

GN HLA-B.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

OC Eutheria; Primates; Catarrhini; Hominidae; Homo.

RN [1]

RA CHANDANAYONG D., SIRIKONG M., LONGTA K., SRINAK D.,

RA SIRIBOONRIT U., RUNGROUNG E., BEJCHANDRA S.;

RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.

DR EMBL: U28759; AAB60367.1;

DR HSSP: P10318; 1ROG.

DR PFAM: PF00129; MHC\_I; 1.

KW MHC.

FT NON\_TER 175 175

FT VARIANT 67 67

FT VARIANT 73 73

FT NON\_TER 175 175

Y -> D.

I -> T.

SQ SEQUENCE 175 AA; 20332 MW; 83A0C5C3 CRC32;

Query Match 50.5%; Score 50; DB 7; Length 175;  
Best Local Similarity 68.8%; Pred. No. 1.3;  
Matches 11; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 TRLNERRENRLALRY 20  
| | | | | | | | | |  
DB 62 TNTQTYRENRLALRY 77

RESULT 11

ID O46703 PRELIMINARY; PRT; 181 AA.  
AC O46703;  
DT 01-JUN-1998 (TREMBLrel. 06, Created)  
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)  
DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)  
DE HLA-B\*51IM (FRAGMENT).  
GN HLA-B\*51IM.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-BLOOD.  
RA GAO X., MATHESON B.;  
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U58320; AAB87725.1; -.  
DR EMBL; U58319; AAB87725.1; JOINED.  
DR HSSP; P10318; IROG.  
DR PFAM; PF00129; MHC\_I; 1.  
KW MHC.  
FT NON\_TER 1  
FT NON\_TER 181  
SQ SEQUENCE 181 AA; 21113 MW; 61FB4281 CRC32;

Query Match 50.5%; Score 50; DB 7; Length 181;  
Best Local Similarity 68.8%; Pred. No. 1.3;  
Matches 11; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 TRLNERRENRLALRY 20  
| | | | | | | | | |  
DB 68 TNTQTYRENRLALRY 83

RESULT 12

ID O62899 PRELIMINARY; PRT; 181 AA.  
AC O62899;  
DT 01-AUG-1998 (TREMBLrel. 07, Created)  
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)  
DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)  
DE MHC CLASS I ANTIGEN (FRAGMENT).  
GN HLA-B.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA MARCOS C.Y., LAZARO A.M., STASTNY P.;  
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF056480; AAC14122.1; -.  
DR EMBL; AF056479; AAC14122.1; JOINED.  
DR HSSP; P10318; IROG.  
DR PFAM; PF00129; MHC\_I; 1.  
KW MHC.  
FT NON\_TER 1  
FT NON\_TER 181  
SQ SEQUENCE 181 AA; 21224 MW; 6736F2EB CRC32;

Query Match 50.5%; Score 50; DB 7; Length 181;  
Best Local Similarity 68.8%; Pred. No. 1.3;  
Matches 11; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 TRLNERRENRLALRY 20  
| | | | | | | | | |  
DB 68 TNTQTYRENRLALRY 83

RESULT 13

ID O62922 PRELIMINARY; PRT; 181 AA.  
AC O62922;  
DT 01-AUG-1998 (TREMBLrel. 07, Created)  
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)  
DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)  
DE MHC CLASS I ANTIGEN (FRAGMENT).  
GN HLA-B.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA STEINER N.K., HURLEY C.K., KOESTER R.P.;  
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF065645; AAC18389.1; -.  
DR EMBL; AF065644; AAC18389.1; JOINED.  
DR HSSP; P10318; IROG.  
DR PFAM; PF00129; MHC\_I; 1.  
KW MHC.  
FT NON\_TER 1  
FT NON\_TER 181  
SQ SEQUENCE 181 AA; 21122 MW; CAE15F67 CRC32;

Query Match 50.5%; Score 50; DB 7; Length 181;  
Best Local Similarity 68.8%; Pred. No. 1.3;  
Matches 11; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 TRLNERRENRLALRY 20  
| | | | | | | | | |  
DB 68 TNTQTYRENRLALRY 83

RESULT 14

ID O19623 PRELIMINARY; PRT; 181 AA.  
AC O19623;  
DT 01-JAN-1998 (TREMBLrel. 05, Created)  
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)  
DE MHC CLASS I ANTIGEN B\*5108 (FRAGMENT).  
GN HLA-B.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA STEINER N.K., NG J., BUSH J., HARTZMAN R.J., JOHNSON-DOW L.,  
HURLEY C.K.;  
RL Hum. Immunol. 56:0-0(1997).  
DR EMBL; U52816; AAB80619.1; -.  
DR EMBL; U52815; AAB80619.1; JOINED.  
DR HSSP; P10318; IROG.  
DR PFAM; PF00129; MHC\_I; 1.  
KW MHC.  
FT NON\_TER 1  
FT NON\_TER 181  
SQ SEQUENCE 181 AA; 21115 MW; A379940B CRC32;

Query Match 50.5%; Score 50; DB 7; Length 181;

Best Local Similarity 68.8%; Pred. No. 1.3;  
Matches 11; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 TRLNERRENRLALRY 20  
| | | | | | | | | |  
Db 68 TNTQTYRENRLALRY 83

RESULT 15  
019747  
ID 019747 PRELIMINARY; PRT; 181 AA.  
AC 019747;  
DT 01-JAN-1998 (Tremblrel. 05, Created)  
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)  
DT 01-NOV-1999 (Tremblrel. 12, Last annotation update)  
DE MHC CLASS I ANTIGEN HLA-B (FRAGMENT).  
GN HLA-B.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA STEINER N.K.; NG J.; BUSH J.; HARTZMAN R.J.; JOHNSON-DOW L.,  
RA HURLEY C.K.;  
RL Hum. Immunol. 56:0-0(1997).  
DR EMBL; U52170; AAB82331.1; -.  
DR EMBL; U52169; AAB82331.1; JOINED.  
DR HSSP; P10318; 1ROG.  
DR -PFAM; PF00129; MHC\_I; 1.  
KW MHC.  
FT NON\_TER 1 1  
FT NON\_TER 181 181  
SQ SEQUENCE 181 AA; 21143 MW; 9CADF6E2 CRC32;

Query Match 50.5%; Score 50; DB 7; Length 181;  
Best Local Similarity 68.8%; Pred. No. 1.3;  
Matches 11; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 TRLNERRENRLALRY 20  
| | | | | | | | | |  
Db 68 TNTQTYRENRLALRY 83

Search completed: February 8, 2000, 19:16:14  
Job time: 21501 sec

**THIS PAGE BLANK (USPTO)**











TITLE	JOURNAL	COMMENT
-------	---------	---------

*	1	566:	contig of 566 bp in length
*			gap of unknown length
*	567	2610:	contig of 2044 bp in length
*			gap of unknown length
*	2611	4919:	contig of 2309 bp in length
*			gap of unknown length
*	4920	7458:	contig of 2539 bp in length
*			gap of unknown length
*	7459	9986:	contig of 2528 bp in length
*			gap of unknown length
*	9987	13374:	contig of 3388 bp in length
*			gap of unknown length
*	13375	16734:	contig of 3360 bp in length
*			gap of unknown length
*	16735	20056:	contig of 3322 bp in length
*			gap of unknown length
*	20057	23962:	contig of 3906 bp in length
*			gap of unknown length
*	23963	28071:	contig of 4109 bp in length
*			gap of unknown length
*	28072	33835:	contig of 5764 bp in length
*			gap of unknown length
*	33836	38938:	contig of 5103 bp in length
*			gap of unknown length
*	38939	44173:	contig of 5235 bp in length
*			gap of unknown length
*	44174	49632:	contig of 5459 bp in length
*			gap of unknown length
*	49633	55995:	contig of 6363 bp in length
*			gap of unknown length
*	55996	62084:	contig of 6089 bp in length
*			gap of unknown length
*	62085	67735:	contig of 5651 bp in length
*			gap of unknown length
*	67736	74775:	contig of 7040 bp in length
*			gap of unknown length
*	74776	82752:	contig of 7977 bp in length
*			gap of unknown length
*	82753	90832:	contig of 8080 bp in length
*			gap of unknown length
*	90833	99824:	contig of 8992 bp in length
*			gap of unknown length
*	99825	110259:	contig of 10435 bp in length
*			gap of unknown length
*	110260	121061:	contig of 10802 bp in length
*			gap of unknown length
*	121062	139035:	contig of 17974 bp in length
*			gap of unknown length
*	139036	181036:	contig of 42001 bp in length
*			gap of unknown length



TITLE Siriboonrit,U., Rungroung,E. and Bejchandra,S.  
JOURNAL B\*51V alleles  
REFERENCE Unpublished  
AUTHORS 2 (bases 1 to 250)  
Chandanayingyong,D., Sirikong,M., Longta,K., Srinak,D.,  
Siriboonrit,U., Rungroung,E. and Bejchandra,S.

TITLE Direct Submission  
JOURNAL Submitted (25-FEB-1997) Transfusion Medicine, Faculty of Medicine,  
Siriraj Hospital, Mahidol University, Prannok Road, Bangkok 10700,  
Thailand

FEATURES Location/Qualifiers  
source  
1..250

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="6"

/map="6p21"  
/haplotype="(a)HLA-A1,B51V,Cw14,DR6(DR52),DQ1/(b)A2,B77,  
Cw8N,DR4(DR53),DQ4"  
/cell\_type="lymphoblastoid"  
/cell\_line="THAI DCH028"

1..250

/gene="HLA-B"  
/note="Allele: HLA-B\*51V; similar to exon 2 of B\*5105"

/number=2  
56 a 82 c 80 g 32 t

BASE COUNT  
ORIGIN

alignment\_scores:

Quality: 50.00 Length: 16  
Ratio: 3.846 Gaps: 0  
Percent Similarity: 81.250 Percent Identity: 68.750

alignment\_block:

US-08-653-294-28 x HSHLABF1 ..

Align seg 1/1 to: HSHLABF1 from: 1 to: 250

5 ThrArgLeuAsnGluArgGluAsnLeuArgIleAlaLeuArgTyr 20  
||||: :||:|||||||  
184 ACCAACACACAGACTTACCGAGAGAACCTCGGATCGCGTCCGCTAC 231

seq\_name: gb\_pr2:HSHLABG1

seq\_documentation\_block:  
LOCUS HSHLABG1 250 bp DNA PRI 25-MAR-1997  
DEFINITION Human cell line THAI DCH011 MHC class I HLA-B gene (allele  
HLA-B\*51V), exon 2.

ACCESSION U90615  
U90615.1 GI:1906037

KEYWORDS

SEGMENT

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE 1 (bases 1 to 250)

AUTHORS Chandanayingyong,D., Sirikong,M., Longta,K., Srinak,D.,

Siriboonrit,U., Rungroung,E. and Bejchandra,S.

TITLE B\*51V alleles

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 250)

AUTHORS Chandanayingyong,D., Sirikong,M., Longta,K., Srinak,D.,

Siriboonrit,U., Rungroung,E. and Bejchandra,S.

TITLE Direct Submission

JOURNAL

Submitted (25-FEB-1997) Transfusion Medicine, Faculty of Medicine,  
Siriraj Hospital, Mahidol University, Prannok Road, Bangkok 10700,  
Thailand

FEATURES Location/Qualifiers

source  
1..250

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/chromosome="6"

/map="6p21"  
/haplotype="(a)HLA-A11.1,B51V,Cw14,DR12(DR52),DQ7/(b)A24,  
B77,Cw8N,DR12(DR52),DQ7"  
/cell\_type="lymphoblastoid"  
/cell\_line="THAI DCH011"  
1..250

exon

/gene="HLA-B"

/note="Allele: HLA-B\*51V; similar to exon 2 of B\*5105"

/number=2

56 a 82 c 80 g 32 t

BASE COUNT  
ORIGIN

alignment\_scores:

Quality: 50.00 Length: 16  
Ratio: 3.846 Gaps: 0  
Percent Similarity: 81.250 Percent Identity: 68.750

alignment\_block:

US-08-653-294-28 x HSHLABG1 ..

Align seg 1/1 to: HSHLABG1 from: 1 to: 250

5 ThrArgLeuAsnGluArgGluAsnLeuArgIleAlaLeuArgTyr 20  
||||: :||:|||||||  
184 ACCAACACACAGACTTACCGAGAGAACCTCGGATCGCGTCCGCTAC 231

seq\_name: gb\_pr2:HSHLABI1

seq\_documentation\_block:

LOCUS HSHLABI1 250 bp DNA PRI 22-MAR-1997  
DEFINITION Human cell line THAI DCH011 MHC class I HLA-B gene (allele  
HLA-B\*1513), exon 2.

ACCESSION U90422

U90422.1 GI:1905834

KEYWORDS

SEGMENT

SOURCE

1 of 2

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE 1 (bases 1 to 250)

AUTHORS Chandanayingyong,D., Sirikong,M., Longta,K., Srinak,D.,

Rungroung,E., Bejchandra,S., Blasczyk,R. and Grosse-Wilde,H.

B15 alleles (B\*1513)

Unpublished

REFERENCE 2 (bases 1 to 250)

AUTHORS Chandanayingyong,D., Sirikong,M., Longta,K., Srinak,D.,

Rungroung,E., Bejchandra,S., Blasczyk,R. and Grosse-Wilde,H.

Direct Submission

TITLE Submitted (23-FEB-1997) Transfusion Medicine, Faculty of Medicine,

Siriraj Hospital, Mahidol University, Prannok Road, Bangkok 10700,

Thailand

FEATURES

Location/Qualifiers

source  
1..250

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/chromosome="VI"

/map="6p21"

/cell\_type="lymphoblastoid"

/cell\_line="THAI DCH011"

1..250

/gene="HLA-B"

/note="Allele: HLA-B\*1513"

/number=2

/product="MHC class I HLA-B"

55 a 83 c 80 g 32 t

BASE COUNT

ORIGIN

alignment\_scores:

Quality: 50.00 Length: 16

Ratio: 3.846 Gaps: 0

2 (ba)

2 (bases 1 to 270)

AUTHORS Poli, F.  
 TITLE Direct Submission  
 JOURNAL Submitted (11-DEC-1996) F. Poli, Centro Trasfusionale e di Immunologie del Trapianti, Ospedale Maggiore Policlinico, Via Francesco Sforza 35, 20122 Milan, ITALY  
 REMARK revised by [3]  
 REFERENCE 3 (bases 1 to 270)  
 AUTHORS Poli, F.  
 JOURNAL Direct Submission  
 TITLE Submitted (03-APR-1997) F. Poli, Centro Trasfusionale e di Immunologie del Trapianti, Ospedale Maggiore Policlinico, Via Francesco Sforza 35, 20122 Milan, ITALY  
 COMMENT On Apr 7, 1997 this sequence version replaced gi:1841917.

## FEATURES

source  
 1..270  
 /organism="Homo sapiens"  
 /isolate="7235"  
 /db\_xref="taxon:9606"  
 /chromosome="6p"  
 /cell\_type="peripheral blood"  
 gene  
 1..270  
 /gene="HLA-B51"  
 exon  
 1..270  
 /gene="HLA-B51"  
 /label=ex2  
 /usedin=V10031:hla\_cds  
 /usedin=V10031:hla\_mrna  
 BASE COUNT 60 a 88 c 84 g 38 t  
 ORIGIN

alignment\_scores:  
 Quality: 50.00 Length: 16  
 Ratio: 3.846 Gaps: 0  
 Percent Similarity: 81.250 Percent Identity: 68.750

## alignment\_block:

US-08-653-294-28 x HSB51EX2 ..

Align seg 1/1 to: HSB51EX2 from: 1 to: 270

5 ThrArgLeuAsnGluArgGluAsnLeuArgIleAlaLeuArgTyr 20  
 ||||| : : : |||||  
 204 ACCAACACACAGACTTACCGAGAGAACCTCGGATCGGCTCCGCTAC 251

seq\_name: gb\_prl:HUMMB512

seq\_documentation\_block:  
 LOCUS HUMMB512 270 bp DNA PRI 07-JAN-1995  
 DEFINITION Human MHC class I HLA-B51 (alpha-1 domain) gene, exon 2.  
 ACCESSION M22787 M21035  
 VERSION M22787.1 GI:187683  
 KEYWORDS HLA-B51 protein; cell surface antigen; cell surface glycoprotein; class I gene; integral membrane protein; major histocompatibility complex.  
 SEGMENT 2 of 7  
 SOURCE Human EBV transformed B lymphoblastoid (LXT-2, Toto cell line) DNA, clone LXT.18, Toto B4.

## ORGANISM

Homo sapiens  
 Eukaryota; Chordata; Craniata; Vertebrata; Mammalia;  
 Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

AUTHORS Hayashi, H., Ennis, P. D., Ariga, H., Salter, R. D., Parham, P., Kano, K. and Takiguchi, M.  
 TITLE HLA-B51 and HLA-Bw52 differ by only two amino acids which are in the helical region of the alpha 1 domain  
 JOURNAL J. Immunol. 142 (1), 306-311 (1989)  
 MEDLINE 85080265

COMMENT Draft entry and computer-readable sequence [1] kindly submitted by P. Ennis, 14-OCT-1988.

## FEATURES

source  
 1..270  
 /organism="Homo sapiens"

/db\_xref="taxon:9606"  
 /map="6p21.3"  
 exon  
 1..270  
 /gene="HLA-B"  
 /note="G00-120-048"  
 /number=2  
 BASE COUNT 60 a 88 c 84 g 38 t  
 ORIGIN Undetermined number of bp after segment 1; chromosome 6p21.3.  
 alignment\_scores:  
 Quality: 50.00 Length: 16  
 Ratio: 3.846 Gaps: 0  
 Percent Similarity: 81.250 Percent Identity: 68.750  
 alignment\_block:  
 US-08-653-294-28 x HUMMB512 ..  
 Align seg 1/1 to: HUMMB512 from: 1 to: 270  
 5 ThrArgLeuAsnGluArgGluAsnLeuArgIleAlaLeuArgTyr 20  
 ||||| : : : |||||  
 204 ACCAACACACAGACTTACCGAGAGAACCTCGGATCGGCTCCGCTAC 251

**THIS PAGE BLANK (USPTO)**



OM of: US-08-653-294-28 to: N\_Geneseq\_36:\* out\_format : pfs  
Date: Feb 8, 2000 7:31 PM  
About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlp  
-O=/cgnl\_1/USPTO\_SPOOL/US08653294/runat\_04022000\_160701\_15807/app\_query.fasta.2  
-DB=N\_Geneseq\_36 -QFMT=fastap -SUFFIX=rng -GAPOP=12.000  
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000  
-GAPOP=4.500 -GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500  
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500  
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62  
-TRANS=human40.cgi -LIST=45 -DOCALLGN=200 -THR\_SCORE=pct  
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0  
-MAXLEN=100000 -USER=US08653294 -NCPU=6 -ICPU=3 -NO\_XLPXY -WAIT  
-THREADS=1

Search information block:  
Query: US-08-653-294-28  
Query length: 20  
Database: N\_Geneseq\_36:\*  
Database sequences: 311585  
Database length: 125096042  
Search time (sec): 873.190000

score_list:	Strd Orig	zScore	EScore	Len	Documentation
N_Geneseq_36:Q29167	+	50.00	154.71	0.5125	270
N_Geneseq_36:Q01834	+	50.00	141.46	2.80	1086
N_Geneseq_36:Q01822	+	50.00	141.46	2.80	1086
N_Geneseq_36:Q05693	+	50.00	141.43	2.81	1089
N_Geneseq_36:Q05701	+	50.00	141.43	2.81	1089
N_Geneseq_36:Q12114	+	50.00	141.43	2.81	1089
N_Geneseq_36:T42608	+	49.00	150.97	0.8283	285
N_Geneseq_36:T60544	+	44.50	134.62	6.74	345
N_Geneseq_36:T60542	+	44.50	132.81	8.50	417
N_Geneseq_36:Q78276	+	43.00	106.51	248.08	3973
N_Geneseq_36:Q78281	+	43.00	106.51	248.08	3973
N_Geneseq_36:Q54386	+	43.00	83.00	5.1e+03	46899
N_Geneseq_36:V65837	+	42.00	116.60	68.01	981
N_Geneseq_36:V49262	+	42.00	105.88	268.84	3023
N_Geneseq_36:Q29114	+	42.00	104.02	341.14	3674
N_Geneseq_36:T75487	+	42.00	103.99	342.61	3687
N_Geneseq_36:T08695	+	42.00	103.62	359.14	3832
N_Geneseq_36:Q90251	+	42.00	100.23	555.13	5474
N_Geneseq_36:V52236	+	42.00	100.08	565.55	5558
N_Geneseq_36:T13265	+	42.00	96.10	941.95	8440
N_Geneseq_36:T07008	+	41.00	135.21	6.25	99
N_Geneseq_36:Q28344	+	41.00	110.37	151.15	1344
N_Geneseq_36:Q30109	+	41.00	99.66	597.17	4140
N_Geneseq_36:V26068	-	41.00	98.12	727.79	4868
N_Geneseq_36:T94023	-	41.00	96.81	860.94	5586
N_Geneseq_36:T14180	+	41.00	91.53	1.7e+03	9718
N_Geneseq_36:T72209	+	41.00	91.25	1.8e+03	10011
N_Geneseq_36:T07447	+	40.00	119.63	46.08	362
N_Geneseq_36:V75032	+	40.00	115.31	80.23	570
N_Geneseq_36:N80745	-	40.00	108.06	203.19	1220
N_Geneseq_36:Q06587	-	40.00	108.06	203.19	1220
N_Geneseq_36:T84985	+	40.00	107.08	230.56	1353
N_Geneseq_36:Q30417	+	40.00	106.97	233.69	1368
N_Geneseq_36:T67839	+	40.00	106.37	252.60	1458
N_Geneseq_36:Q03536	-	40.00	106.10	261.30	1499
N_Geneseq_36:N71241	-	40.00	106.10	261.51	1500
N_Geneseq_36:N70058	-	40.00	106.10	261.51	1500
N_Geneseq_36:Q90042	-	40.00	106.10	261.51	1500
N_Geneseq_36:T60306	-	40.00	106.10	261.51	1500
N_Geneseq_36:N60475	-	40.00	105.94	266.63	1524

N\_Geneseq\_36:T84986 + 40.00 101.78 455.07 2361 ! Burkholderia cepacia insert  
N\_Geneseq\_36:X25097 + 40.00 100.29 550.67 2760 ! Bacillus subtilis metallopro  
N\_Geneseq\_36:N82261 + 40.00 98.75 670.53 3243 ! Enkephalinase gene (rat). D  
N\_Geneseq\_36:Q62679 + 40.00 98.75 670.53 3243 ! Rat enkephalinase cDNA. Met

seq\_name: N\_Geneseq\_36:Q29167

seq\_documentation\_block:

ID Q29167 standard; DNA; 270 BP.

AC Q29167:

DT 09-MAR-1993 (first entry)

DE HLA-Bw 52 exon 2 alpha-1 domain.

KW Human leukocyte antigen; transgenic; germ cells; somatic cells;

KW expression; ss.

PN J04091731-A.

PD 25-MAR-1992.

PF 03-AUG-1990; 207329.

PR 03-AUG-1990; JPM-207329.

PA (OLYU ) OLYMPUS OPTICAL CO.

DR WPI; 92-342893/42.

PT Transgenic non-human mammalian HLA-Bw 52 gene - useful for

PT analysis of expression of gene structure, and prodn. of

PT mouse model of human disease

PS Disclosure; Fig 1; 8pp; Japanese.

CC The sequence shows the exon 2 alpha-1-domain of the human leukocyte

CC antigen-Bw 52 gene. The complete gene may be introduced into non-

CC human mammals, pref. rat or mouse, or their ancestors at the primary

CC developmental biological step via transplantation into the zygote or

CC embryo to generate transgenic non-human mammals incorporating the

CC HLA-Bw 52 gene in both their germ cells and somatic cells. Transgenic

CC non-human mammals contg. HLA-Bw 52 are useful for the analysis of

CC expression of the gene, its structure, and prodn. of mouse models of

CC human disease. See also Q29166-72.

SQ Sequence 270 BP; 59 A; 88 C; 86 G; 37 T;

alignment\_scores:

Quality: 50.00 Length: 16

Ratio: 3.846 Gaps: 0

Percent Similarity: 81.250 Percent Identity: 68.750

alignment\_block:

US-08-653-294-28 x Q29167 ..

Align seg 1/1 to: Q29167 from: 1 to: 270

5 ThrArgLeuAsnGluArgArgGluAsnLeuArgIleAlaLeuArgTyr 20

||||: ||: ||||| ||||| ||||| ||||| ||||| |||||

204 ACCAACACACAGACTTACCGAGAGACCTCGGATCGGCTCGCTAC 251

seq\_name: N\_Geneseq\_36:Q01834

seq\_documentation\_block:

ID Q01834 standard; DNA; 1086 BP.

AC Q01834:

DT 19-MAR-1991 (first entry)

DE Sequence encoding HLA-B51 antigen.

KW Probe: HLA class I DNA; immunogen; ss.

OS Homo sapiens.

PN EP354580-A.

PD 14-FEB-1990.

PF 10-AUG-1989.

PR 11-AUG-1988; JP-200758.

PA (OLYU) Olympus Optical Co., Ltd.

PI Kano K, Takiguchi;

DR WPI; 90-046289/07.

PT New DNA for class 1 human leukocyte antigens and derived probes and

PT transformed cells, useful for DNA typing, as immunogens etc.

PS Claim 1; Page 11; 23pp; English.

CC The HLA class I DNA can be used as a source of probes for use in DNA

CC typing. Transformed cells, which are useful as immunogens, can be

CC obtained by introducing these DNAs into eucaryotic cells.

SQ Sequence 1086 BP; 224 A; 334 C; 356 G; 172 T;

```

alignment_scores:
  Quality: 50.00      Length: 16
  Ratio: 3.846       Gaps: 0
  Percent Similarity: 81.250  Percent Identity: 68.750

alignment_block:
US-08-653-294-28 x Q01834 ..
  Align seg 1/1 to: Q01834 from: 1 to: 1086

  5 ThrArgLeuAsnGluArgGluAsnLeuArgIleAlaLeuArgTyr 20
  ||||: :: |||||
  276 ACCACACACACACTTACCGAGAGAACCTGCGGATCGCGCTCCGTAC 323

seq_name: N_Geneseq_36:Q01822

seq_documentation_block:
ID Q01822 standard; DNA; 1086 BP.
AC Q01822;
DT 19-MAY-1991 (first entry)
DE Sequence encoding HLA-Bw52 antigen.
KW Probe: HLA class I DNA; immunogen; ss.
OS Homo sapiens.
FH key Location/Qualifiers
FT cds 1..1086
FT /*tag= a
PN EP-354580-A.
PD 14-FEB-1990.
PF 10-AUG-1989.
PR 11-AUG-1988; JP-200758.
PA (OLYU) Olympus Optical Co., Ltd.
PI Kano K. Takiguchi;
DR WPI; 90-046289/07.
DR P-PSDB; R03142.
PT New DNA for class I human leucocyte antigens and derived probes and
PT transformed cells, useful for DNA typing, as immunogens etc.
PS Claim 2; pp11-12; 23pp; English.
CC The HLA class I DNA can be used as a source of probes for use in DNA
CC typing. Transformed cells, which are useful as immunogens, can be
CC obtained by introducing these DNAs into eucaryotic cells.
SQ Sequence 1086 BP; 223 A; 335 C; 358 G; 170 T;

alignment_scores:
  Quality: 50.00      Length: 16
  Ratio: 3.846       Gaps: 0
  Percent Similarity: 81.250  Percent Identity: 68.750

alignment_block:
US-08-653-294-28 x Q01822 ..
  Align seg 1/1 to: Q01822 from: 1 to: 1086

  5 ThrArgLeuAsnGluArgGluAsnLeuArgIleAlaLeuArgTyr 20
  ||||: :: |||||
  276 ACCACACACACACTTACCGAGAGAACCTGCGGATCGCGCTCCGTAC 323

seq_name: N_Geneseq_36:Q05693

seq_documentation_block:
ID Q05693 standard; DNA; 1089 BP.
AC Q05693;
DT 03-JAN-1991 (first entry)
DE HLA-B51 gene for production of monoclonal antibodies.
KW Allotype specific monoclonal anti-HLA antibodies; hybridomas;
KW transgenic animals; HLA-B51 gene; ss.
FH key Location/Qualifiers
FT exon 1..73
FT /*tag= a
FT /*number=1
FT /*tag= b
FT /*number=2

alignment_scores:
  Quality: 50.00      Length: 16
  Ratio: 3.846       Gaps: 0
  Percent Similarity: 81.250  Percent Identity: 68.750

alignment_block:
US-08-653-294-28 x Q05693 ..
  Align seg 1/1 to: Q05693 from: 1 to: 1089

  5 ThrArgLeuAsnGluArgGluAsnLeuArgIleAlaLeuArgTyr 20
  ||||: :: |||||
  277 ACCACACACACACTTACCGAGAGAACCTGCGGATCGCGCTCCGTAC 324

seq_name: N_Geneseq_36:Q05701

seq_documentation_block:
ID Q05701 standard; DNA; 1089 BP.
AC Q05701;
DT 03-JAN-1991 (first entry)
DE HLA-Bw52 gene for production of monoclonal antibodies.
KW Allotype specific monoclonal anti-HLA antibodies; hybridomas;
KW transgenic animals; HLA-Bw52 gene; ss.
FH key Location/Qualifiers
FT exon 1..73
FT /*tag= a
FT /*number=1
FT /*tag= b
FT /*number=2

FT /*tag= b
FT /*number=2
FT /*note="alpha 1-domain"
FT 344..619
FT /*tag= c
FT /*number=3
FT /*note="alpha 2-domain"
FT 620..895
FT /*tag= d
FT /*number=4
FT /*note="alpha 3-domain"
FT 896..1012
FT /*tag= e
FT /*number=5
FT 1013..1042
FT /*tag= f
FT /*number=6
FT 1043..1089
FT /*tag= g
FT /*number=7
PN EP-383183-A.
PD 22-AUG-1990.
PF 07-FEB-1990; 102424.
PR 08-FEB-1989; JP-029313.
PA (OLYU) OLYMPUS OPTICAL KK.
PI Takiguchi M;
DR WPI; 90-255479/34.
PT Allotype specific monoclonal anti- HLA antibodies prodn. - using
PT hybridomas derived from transgenic animals carrying HLA gene and
PT immunised with HLA antigen of different allotype
PS Disclosure; Fig 1 A-G; 20pp; English.
CC The human HLA-B51 gene was injected into fertilised mouse eggs and
CC then these introduced into the uterus of a pseudo pregnant mouse.
CC The young were tested to ensure incorporation of the gene into the
CC chromosome, and one of them mated 3 times with a normal male to
CC produce 16 young, seven of which carried the HLA-B51 gene.
CC The transgenic offspring were immunised with HLA antigen.
CC The spleen lymphocytes were fused with myeloma cells. Hybridomas
CC producing antibodies were selected.
CC See also Q05701.
SQ Sequence 1089 BP; 224 A; 335 C; 357 G; 173 T;

alignment_scores:
  Quality: 50.00      Length: 16
  Ratio: 3.846       Gaps: 0
  Percent Similarity: 81.250  Percent Identity: 68.750

alignment_block:
US-08-653-294-28 x Q05693 ..
  Align seg 1/1 to: Q05693 from: 1 to: 1089

  5 ThrArgLeuAsnGluArgGluAsnLeuArgIleAlaLeuArgTyr 20
  ||||: :: |||||
  277 ACCACACACACACTTACCGAGAGAACCTGCGGATCGCGCTCCGTAC 324

seq_name: N_Geneseq_36:Q05701

seq_documentation_block:
ID Q05701 standard; DNA; 1089 BP.
AC Q05701;
DT 03-JAN-1991 (first entry)
DE HLA-Bw52 gene for production of monoclonal antibodies.
KW Allotype specific monoclonal anti-HLA antibodies; hybridomas;
KW transgenic animals; HLA-Bw52 gene; ss.
FH key Location/Qualifiers
FT exon 1..73
FT /*tag= a
FT /*number=1
FT /*tag= b
FT /*number=2

```

```

FT exon /note="alpha 1-domain"
FT 344..619
FT /*tag= c
FT /number=3
FT /note="alpha 2-domain"
FT 620..895
FT /*tag= d
FT /number=4
FT /note="alpha 3-domain"
FT 896..1012
FT /*tag= e
FT /number=5
FT 1013..1042
FT /*tag= f
FT /number=6
FT 1043..1089
FT /*tag= g
FT /number=7
FT
FT EP-383183-A.
FT 22-AUG-1990.
FT 07-FEB-1990; 102424.
FT 08-FEB-1989; JP-029313.
FT (OLYU ) OLYMPUS OPTICAL KK.
FT Takiguchi M;
FT WPI; 90-255479/34.
FT Allotype specific monoclonal anti- HLA antibodies prodn. - using
FT hybridomas derived from transgenic animals carrying HLA gene and
FT immunised with HLA antigen of different allotype
FT Disclosure; Fig 1 A-G; 20pp; English.
FT The human HLA-Bw52 gene was introduced into mouse L cells and
FT then these cells used to immunise one of the transgenic mice
FT (See Q05693).
FT The spleen lymphocytes were fused with myeloma cells (P3x63-Ag8.653).
FT Hybridomas producing antibodies were selected.
FT Sequence 1089 BP; 223 A; 336 C; 359 G; 171 T;

alignment_scores:
Quality: 50.00 Length: 16
Ratio: 3.846 Gaps: 0
Percent Similarity: 81.250 Percent Identity: 68.750

alignment_block:
US-08-653-294-28 x Q05701 ..
Align seg 1/1 to: Q05701 from: 1 to: 1089

5 ThrArgLeuAsnGluArgGluAsnLeuArgIleAlaLeuArgTyr 20
||||: |||
277 ACCAACACACAGACTTACCGAGAGAACCTGCGGATCGCGCTCCGCTAC 324

seq_name: N_Geneseq_36:Q12114

seq_documentation_block:
ID Q12114 standard; DNA; 1089 BP.
AC Q12114;
DT 29-AUG-1991 (first entry)
DE HLA-Bw53 exon.
KW Human leukocyte antigen; probe; major histocompatibility complex;
KW MHC; class I; ss.
OS Homo sapiens.
FH key Location/Qualifiers
FT cds 1..1089
FT /*tag= a
FT J03112487-A.
FT 14-MAR-1991.
FT 22-SEP-1989; 247697.
FT 22-SEP-1989; JP-247697.
FT (OLYU ) OLYMPUS OPTICAL KK.
FT WPI; 91-182991/25.
FT P-PSDB; R12483.
FT HLA-Bw53 gene, DNA probe and transformant cells - used for
FT immunisation, identifying specificity of antiserum etc.

```

```

PS Claim 1; Page 1; lipp; Japanese.
CC Probes comprising part of the sequence can be used to identify
CC Class I genes. The DNA can be expressed for immunisation of
CC animals and prodn. of monoclonal antibodies specific for the
CC HLA-Bw53 antigen. See also J03112485 and J03112486. 174 T;
SQ Sequence 1089 BP; 222 A; 337 C; 356 G; 174 T;

alignment_scores:
Quality: 50.00 Length: 16
Ratio: 3.846 Gaps: 0
Percent Similarity: 81.250 Percent Identity: 68.750

alignment_block:
US-08-653-294-28 x Q12114 ..
Align seg 1/1 to: Q12114 from: 1 to: 1089

5 ThrArgLeuAsnGluArgGluAsnLeuArgIleAlaLeuArgTyr 20
||||: |||
277 ACCAACACACAGACTTACCGAGAGAACCTGCGGATCGCGCTCCGCTAC 324

seq_name: N_Geneseq_36:T42608

seq_documentation_block:
ID T42608 standard; DNA; 285 BP.
AC T42608;
DT 07-AUG-1997 (first entry)
DE Synthetic gene, LqhIV encoding scorpion neurotoxin.
KW neurotoxin; pest control; insect; acarid; nematode; LqhIV; LqhVI;
KW venom; Leirus quinquatratrius hebraeus; scorpion; lepidoptera;
KW baculovirus; arthropod; ds.
OS Synthetic.
FH key Location/Qualifiers
FT cds 70..267
FT /*tag= a
FT /note= "no start codon given"
FT W09636221-A2.
FT 21-NOV-1996.
FT 30-APR-1996; U06076.
FT 08-MAY-1995; US-435040.
FT (REGC ) UNIV CALIFORNIA.
FT Hammock BD, Herrmann R, Moskowitz H;
FT WPI; 97-011744/01.
FT P-PSDB; W06339.
FT Control of insects, acarids and nematodes - using novel scorpion
FT toxins or a combination of 2 or more insect toxins
PS Claim 1; Fig 1; 35pp; English.
CC A novel method for controlling pests selected from insects, acarids, and
CC nematodes, comprises treating the pests or their loci with at least two
CC different insect toxins, the source of the toxins being at least one
CC recombinant microbe, the toxins having non-overlapping binding sites at
CC an insect cellular membrane channel. The present sequence is a synthetic
CC gene for LqhIV toxin, derived from the venom of Leirus quinquatratrius
CC hebraeus (Scorpion). LqhIV is the most potent lepidopterous toxin
CC isolated from scorpion venom to date, whilst LqhVI (W06340) has weak
CC mammal toxicity. In particular the insect toxin genes are expressed
CC by baculovirus vectors, as the baculoviruses infect only arthropods,
CC therefore pose little or no risk to humans, plants or the environment.
SQ Sequence 285 BP; 67 A; 92 C; 70 G; 56 T;

alignment_scores:
Quality: 49.00 Length: 19
Ratio: 3.062 Gaps: 0
Percent Similarity: 84.211 Percent Identity: 52.632

alignment_block:
US-08-653-294-28 x T42608 ..
Align seg 1/1 to: T42608 from: 1 to: 285

1 TyrArgLeuAlaThrArgLeuAsnGluArgGluAsnLeuArgIleAl 17

```

```

:::||||| ||| |||||:::|||||:::|||||:::
66 CACCGCGTGGCGAGCGCTACATCGCGACACGAAGACTGCGGTACA 115
17 aLeuArg 19
|||||
116 CACTCGCG 122

```

seq\_name: N\_Geneseq\_36:T60544

seq\_documentation\_block:

```

ID T60544 standard; DNA; 345 BP.
AC T60544;
DE 28-NOV-1997 (first entry)
DE (Pro23-Lys37delAsn37-Ile134) truncated GDNF DNA.
KW Glial cell line-derived neurotrophic factor; GDNF; human;
KW dopaminergic; nerve cell; Parkinson's disease; gene therapy; ss.
OS Chimeric Homo sapiens;
OS Chimeric synthetic.
PN WO9711964-A1.
PD 03-APR-1997.
PF 16-SEP-1996; U14915.
PR 28-SEP-1995; US-535681.
PA (AMGE-) AMGEN INC.
PI Hu SS;
DR WPI; 97-212849/19.
DR P-PSDB; W15743.
PT Truncated glial cell line-derived neurotrophic factor protein - used
PT in the treatment and gene therapy of Parkinson's disease
PS Claim 17; Page 75-76; 105pp; English.
CC This polynucleotide encodes (Pro23-Lys37delAsn37-Ile134)-truncated
CC glial cell line-derived neurotrophic factor (GDNF) protein
CC (W15743). It was constructed using recombinant genetic engineering
CC techniques and can be used to produce recombinant truncated GDNF in
CC transformed or transfected host (esp. E. coli or CHO) cells. Also
CC claimed are a vector, the host cell and a composition comprising
CC mature (see W15706) and truncated GDNFs. The truncated GDNF can be
CC used in the treatment of nervous system damage, esp. Parkinson's
CC disease (claimed). The polynucleotide can be used to provide in
CC vivo production of truncated GDNF, thereby providing gene therapy.
SQ Sequence 345 BP; 79 A; 101 C; 79 G; 86 T;

```

alignment\_scores:

Quality:	44.50	Length:	21
Ratio:	2.781	Gaps:	1
Percent Similarity:	76.190	Percent Identity:	52.381

alignment\_block:

US-08-653-294-28 x T60544 ..

Align seg 1/1 to: T60544 from: 1 to: 345

```

1 TyrArgLeuAlaThrArgLeu...AsnGluArgGluAsnLeuArgII 16
||||| :::::||||| ||:::||||| ||||| ::
90 TACCGACCTGGGTCTCGGTTACGAACCAAGAAAGAAATTAATCTCGGTT 139
16 eAlaLeuArgTyr 20
: |||||:::
140 ACTGCTCCGGTTC 152

```

seq\_name: N\_Geneseq\_36:T60542

seq\_documentation\_block:

```

ID T60542 standard; DNA; 417 BP.
AC T60542;
DE 28-NOV-1997 (first entry)
DE MetGDNF degenerate DNA sequence.
KW Glial cell line-derived neurotrophic factor; GDNF; human;
KW dopaminergic; nerve cell; Parkinson's disease; gene therapy; ss.
OS Chimeric Homo sapiens;
OS Chimeric synthetic.
FH Key Location/Qualifiers
FT mat_peptide 4..439

```

FT WO9711964-A1. /\*tag= a

```

PN 03-APR-1997.
PD 16-SEP-1996; U14915.
PR 28-SEP-1995; US-535681.
PA (AMGE-) AMGEN INC.
PI Hu SS;
DR WPI; 97-212849/19.
DR Truncated glial cell line-derived neurotrophic factor protein - used
DR in the treatment and gene therapy of Parkinson's disease
PS Claim 15; Page 74-75; 105pp; English.
CC This DNA sequence represents an alternative sequence encoding
CC N-terminally methionylated human mature glial cell line-derived
CC neurotrophic factor (GDNF) protein (see W15706). It can be used
CC for the production of claimed truncated GDNF proteins in
CC transformed or transfected host (esp. E. coli or CHO) cells.
CC Truncated GDNF proteins have the formula: X-(Cys41-Cys133)-Y, where
CC (Cys41-Cys133) = Cys41 through Cys133 of mature human GDNF; Y = a
CC C-terminal Cys133 or Ile134; and X = a methionylated or
CC nonmethionylated amine group of Cys41 or an amino terminus sequence
CC selected from G, RG, NRG or the sequences given in W15707-42.
CC Also claimed are: the polynucleotides (see T60542-46) encoding
CC truncated GDNF (see W15743-45); a vector; the host cells; and a
CC GDNF composition comprising mature and truncated GDNFs useful in
CC the treatment of nervous system damage, esp. Parkinson's disease
CC (claimed). The polynucleotides may be used to provide in vivo
CC production of truncated GDNF, thereby providing gene therapy.
SQ Sequence 417 BP; 95 A; 123 C; 99 G; 100 T;

```

alignment\_scores:

Quality:	44.50	Length:	21
Ratio:	2.781	Gaps:	1
Percent Similarity:	76.190	Percent Identity:	52.381

alignment\_block:

US-08-653-294-28 x T60542 ..

Align seg 1/1 to: T60542 from: 1 to: 417

```

1 TyrArgLeuAlaThrArgLeu...AsnGluArgGluAsnLeuArgII 16
||||| :::::||||| ||:::||||| ||||| ::
156 TACCGACCTGGGTCTCGGTTACGAACCAAGAAAGAAATTAATCTCGGTT 205
16 eAlaLeuArgTyr 20
: |||||:::
206 ACTGCTCCGGTTC 218

```

seq\_name: N\_Geneseq\_36:Q78276

seq\_documentation\_block:

```

ID Q78276 standard; DNA; 3973 BP.
AC Q78276;
DE 06-JUN-1995 (first entry)
DE T. niveum alanine-racemase.
KW Alanine-racemase; glyceraldehyde-3-phosphate-dehydrogenase;
KW cyclosporin; immunosuppressive; ss.
OS Tolypocladium niveum ATCC 34921.
PN WO9425606-A.
PD 10-NOV-1994.
PF 23-APR-1994; E01272.
PR 23-APR-1993; DE-312856.
PR 04-MAY-1993; DE-314610.
PR 17-MAY-1993; DE-316419.
PA (SANO ) SANDOZ LTD.
PA (SANO ) SANDOZ PATENT GMBH.
PA (SANO ) SANDOZ-ERFINDUNGS VERW GES MBH.
PI Kocher HP, Schneider-scherzer E, Schoergendorfer K;
PI Weber G;
DR WPI; 94-358281/44.
PT New nucleic acid encoding eukaryotic alanine racemase - and
PT related vectors, host cells and recombinant enzyme, useful for
PT producing cyclosporin derivs. or increasing cyclosporin prodn.,

```





OS Synthetic.  
 FH Key Location/Qualifiers  
 FT misc\_feature 652..654  
 FT /\*tag= a  
 FT /label= Initiation\_codon  
 PN WO9215672-A.  
 PD 17-SEP-1992.  
 PF 09-MAR-1992; U01906.  
 PR 07-MAR-1991; US-666056.  
 PR 11-JUN-1991; US-713967.  
 PR 06-MAR-1992; US-847951.  
 PA (VIR-) VIROGENETICS CORP.  
 PI Cox WI, De Taisne C, Francis J, Gettig RR, Johnson GP,  
 PI Limbach KJ, Norton EK, Paolletti E, Perkins ME, Pincus SE,  
 PI Riviere M, Tartaglia J, Taylor J;  
 DR WPI: 92-331718/40.  
 PT Vaccine comprises recombinant, attenuated pox-virus - use for  
 PT vaccinating against viral infections such as rabies, hepatitis B,  
 PT HIV, HSV, EBV, CMV, mumps etc.  
 PS Disclosure: Fig 27; 456pp; English.  
 CC The sequence given encodes the feline leukemia virus (FeLV-A) gag  
 CC gene. This sequence was used in the construction of an ALVAC-based  
 CC recombinant containing the FeLV-A gag gene. The gag/pol sequence was  
 CC derived from plasmid pFGA-2 gag. The gag gene was liberated on a 2.5  
 CC kb fragment and was amplified by polymerase chain reaction (PCR).  
 CC This sequence was aligned with the vaccinia virus H6 promoter. The  
 CC FeLV-A pol gene was also included in this construction to yield a  
 CC plasmid designated PC3FCGAG. Termination codons were included to form  
 CC the plasmid PC3FCGAGVQ. The FeLV gag/pol expression cassette was  
 CC excised from this plasmid and designated PC3DORAGVQ. PC3DORAGVQ  
 CC was used in standard in vitro recombination assays with vCP83 and  
 CC vCP87 as rescue viruses. Recombinants containing the entire FeLV-A  
 CC gag/pol sequences and the entire FeLV-A env gene were designated  
 CC vCP97, while recombinants containing the same gag/pol sequences and  
 CC the entire FeLV-A env lacking the immuno-suppressive region were  
 CC designated vCP93. The FeLV-A gag can be inserted into a vaccinia  
 CC virus background by including a polylinker sequence flanking the  
 CC coding region and using these sequences to insert the gag region into  
 CC the A type inclusion body region (ATI) of the plasmid pSB541. See  
 CC also Q35501-864.  
 SQ Sequence 3674 BP; 970 A; 1018 C; 883 G; 803 T;

## alignment\_scores:

Quality: 42.00 Length: 11  
 Ratio: 3.818 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 72.727

## alignment\_block:

US-08-653-294-28 x Q29114 ..

Align seg 1/1 to: Q29114 from: 1 to: 3674

3 LeuAlaThrArgLeuAsnGluArgGluAsn 13  
 ::::|||||:|||||:|||||:|||||:|||||:  
 1408 ATTGCAAGCGCGCTAAGGGAACGACGAGAAAC 1440

**THIS PAGE BLANK (USPTO)**



Strid	Orig	ZScore	EScore	Len	Documentation
gb_est1:D35316	+	52.00	154.70	360	D35316 CELK019G2F Yuji Kohara u
gb_est1:D35321	+	52.00	154.70	360	D35321 CELK019G6F Yuji Kohara u
gb_est1:D35579	+	52.00	154.70	360	D35579 CELK023G9F Yuji Kohara u
gb_est17:C69731	+	52.00	154.70	360	C69731 Yuji Kohara unpub
gb_est17:C69775	+	52.00	154.70	360	C69775 C69775 Yuji Kohara unpub
gb_est17:C69958	+	52.00	154.70	360	C69958 C69958 Yuji Kohara unpub
gb_est1:D37592	+	52.00	153.72	399	D37592 CELK004CAF Yuji Kohara u
gb_est1:C03945	+	50.00	152.61	232	C03945 C3945 Human heart CDNA
gb_est11:A2863158	+	50.00	150.72	12.51	A2863158 PMY0534 KGI-a Lambda Z
gb_est17:D82221	+	50.00	148.04	375	D82221 HUMHG6426 Human pancrea
gb_est31:A1936864	+	50.00	141.48	40.91	A1936864 wC74h11.x1 NCICGAPPa
gb_est10:AA151891	+	49.00	148.57	255	AA151891 z001f06.r1 Strategene
gb_est36:AV186045	+	49.00	145.30	25.47	AV186045 AV186046.r1 Kohara u
gb_est33:A1777709	+	49.00	144.46	27.90	A1777709 EST358504 tomato susce
gb_est33:A1774937	+	49.00	142.25	37.06	A1774937 EST256037 x01 NCICGAPB
gb_est36:A1359260	+	49.00	140.16	48.45	A1359260 qy27b07.x1 NCICGAPB
gb_gss6:AQ849043	+	48.00	139.27	54.31	AQ849043 F2F8Fr TGF Arabidopsis
gb_gss3:C82809	+	48.00	139.11	55.39	C82809 HS_5138_B2_B06 SP6E RP
gb_gss15:AQ651593	+	47.00	136.17	80.77	AQ651593 HS_5138_B2_B06 SP6E RP
gb_gss8:AQ152690	+	47.00	131.32	150.58	AQ152690 HS_2231.A1.A08 MF C1T
gb_est41:AV337157	+	46.50	141.68	39.67	AV337157 AV317157 RIKEN full-le
gb_est17:C699258	+	46.00	135.89	83.74	C699258 C69258 Yuji Kohara unpub
gb_gss6:AQ846741	+	46.00	135.06	93.19	AQ846741 FMAJFV1.lm45f01.x1 Leis
gb_est39:AW100149	+	46.00	134.54	99.59	AW100149 s226a03.y1 Gm-cl012 GL
gb_est31:A1676428	+	46.00	132.79	124.69	A1676428 etmEST0183 ERI Biomeri
gb_gss15:AQ693547	+	46.00	130.14	175.02	AQ693547 HS_3462_A2_C02_SP6E RP
gb_gss15:AQ849574	+	46.00	129.60	187.73	AQ849574 Sheared DNA-28D09.TF SH
gb_gss6:AQ853602	+	46.00	128.41	128.66	AQ853602 CpG1695A CPIOAGNAD1
gb_gss1:CNS00A0A	+	46.00	127.46	246.95	CNS00A0A Drosophila melanogaste
gb_gss1:CNS0043C	+	46.00	126.84	267.42	CNS0043C Drosophila melanogaste
gb_est43:AV352495	+	45.50	140.06	49.09	AV352495 AV352495 RIKEN full-le
gb_gss6:AQ046854	+	45.00	142.97	33.79	AQ046854 RVC111-33120.TK RPCI-1
gb_gss6:AQ876554	+	45.00	131.76	142.33	AQ876554 HS_2082_B2.A07_TTC C1T
gb_est23:A1124815	+	45.00	131.43	148.43	A1124815 A56e06.x1 Johnston fty
gb_gss15:AQ640208	+	45.00	130.18	174.16	AQ640208 927P1.TF 927P1.TF
gb_gss4:AQ720034	+	45.00	129.05	201.51	AQ720034 HS_5541.A2.G07.SP6E RP
gb_gss9:AQ140736	+	45.00	128.66	211.72	AQ140736 HS_3124_A2.D09.T7 C1T
gb_est10:AA147151	+	45.00	128.21	224.37	AA147151 z032d06.r1 Strategene
gb_gss15:AQ655562	+	45.00	127.48	246.21	AQ655562 Sheared DNA-9C11.TF SH
gb_gss7:AQ0900542	+	45.00	122.47	468.11	AQ0900542 HS_5141.A2.B01.T7A RP
gb_est13:A4359768	+	44.00	132.61	127.60	A4359768 EST68859 Fetal lung IT
gb_est7:W91371	+	44.00	131.52	146.77	W91371 MTA.H10J079.B MTA adult

```

Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis
1 (bases 1 to 360)
Kohara,Y., Mitsuki,H., Nishigaki,A., Motohashi,T., Sugimoto,A. and
Tabara,H.
TITLE      Toward an expression map of the C.elegans genome
JOURNAL    Unpublished (1994)
COMMENT    Contact: Yuji Kohara
           Gene Library Lab
           National Institute of Genetics
           Yata 1111, Mishima, Shizuoka 411, Japan
           Tel: 81-559-81-6854
           Fax: 81-559-81-6855
           Email: ykohara@lab.nig.ac.jp.
           Location/Qualifiers
             1..360
               /organism="Caenorhabditis elegans"
               /strain="CB1489 him-8(e1489)"
               /db_xref="taxon:6239"
               /clone="yk19g6"
               /clone_lib="Yuji Kohara unpublished cDNA"
               /sex="hermaphrodite, male"
               /tissue_type="whole animal"
               /dev_stage="varied"
               /dev_stage="varied"
BASE COUNT 116 a      85 c      79 g      80 t
ORIGIN

alignment_scores:
  Quality: 52.00      Length: 17
  Ratio: 3.714      Gaps: 0
  Percent Similarity: 82.353      Percent identity: 70.588

alignment_block:
US-08-653-294-28 x D35321 ..

Align seg 1/1 to: D35321 from: 1 to: 360

2  ArgLeuAlaThrArgLeuAsnGluArgGluAsnLeuArgIleAlaLe 18
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
79 AGATCCGGCTTAAGCTTGACGGAGAGAGATCCATCTGAGGATTATCT 128

18 u 18
   |
129 G 129

seq_name: gb_estl:D35579

seq_documentation_block:
LOCUS      D35579          360 bp      mRNA          EST          08-AUG-1994
DEFINITION CELK02269F Yuji Kohara unpublished cDNA Caenorhabditis elegans cDNA
clone yk22g 5', mRNA sequence.
ACCESSION  D35579
VERSION    D35579.1  GI:527030
KEYWORDS   EST.
SOURCE     Caenorhabditis elegans.
           Caenorhabditis elegans.
           Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
           Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
1 (bases 1 to 360)
Kohara,Y., Mitsuki,H., Nishigaki,A., Motohashi,T., Sugimoto,A. and
Tabara,H.
TITLE      Toward an expression map of the C.elegans genome
JOURNAL    Unpublished (1994)
COMMENT    Contact: Yuji Kohara
           Gene Library Lab
           National Institute of Genetics
           Yata 1111, Mishima, Shizuoka 411, Japan
           Tel: 81-559-81-6854
           Fax: 81-559-81-6855
           Email: ykohara@lab.nig.ac.jp.
           Location/Qualifiers
             1..360
               /organism="Caenorhabditis elegans"

```

seq\_name: gb\_est17:C69958

```

JOURNAL      Unpublished (1994)
COMMENT      Contact: Yuji Kohara
              Gene Library Lab
              National Institute of Genetics
              Yata 1111, Mishima, Shizuoka 411, Japan
              Tel: 81-559-81-6854
              Fax: 81-559-81-6855
              Email: ykohara@lab.nig.ac.jp.

FEATURES     Location/Qualifiers
            1..399
            /organism="Caenorhabditis elegans"
            /strain="CB1489 him-8(e1489)"
            /db_xref="taxon:6239"
            /clone="yk4c4"
            /clone_lib="Yuji Kohara unpublished cDNA"
            /sex="hermaphrodite, male"
            /tissue-type="whole animal"
            /dev_stage="varied"

BASE COUNT  124 a 93 c 95 g 86 t 1 others
ORIGIN

alignment_scores:
  Quality: 52.00      Length: 17
  Ratio: 3.714       Gaps: 0
  Percent Similarity: 82.353 Percent Identity: 70.588

alignment_block:
US-08-653-294-28 x D37592 ..
Align seg 1/1 to: D37592 from: 1 to: 399

2 ArgLeuAlaThrArgLeuAsnGluArgGluAsnLeuArgIleAlaLe 18
|||||
74 AGATCCCGGTAGGTTCACGGAGAGAGATCCCACTGAGGATTATCT 123

18 u 18
|
124 G 124

seq_name: gb_est8:C03945

seq_documentation_block: 232 bp mRNA EST 30-JUL-1996
LOCUS C03945 Human heart cDNA (Ynakamura) Homo sapiens cDNA clone
DEFINITION 3NHC2454, mRNA sequence.
ACCESSION C03945
VERSION C03945.1 GI:1467196
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
Tanaka.T., Oglwara.A., Uchiyama.I., Takagi.T., Yazaki.Y. and
Nakamura.Y.
REFERENCE 1 (bases 1 to 232)
AUTHORS Construction of a normalized directionally cloned cDNA library from
adult heart and analysis of 3040 clones by partial sequencing
Genomics 35 (1), 231-235 (1996)
JOURNAL 96299782
MEDLINE
COMMENT On Oct 24, 1995 this sequence version replaced gi:1040105.
Contact: Yusuke Nakamura
Institute of Medical Science
University of Tokyo
4-6-1, Shirokanedai, Minato-ku, Tokyo 108, Japan
Tel: 81-3-5449-5372
Fax: 81-3-5449-5433
Email: yusuke@ims.u-tokyo.ac.jp.
Location/Qualifiers
1..232
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="3NHC2454"

JOURNAL      Unpublished (1994)
COMMENT      Contact: Yuji Kohara
              Gene Library Lab
              National Institute of Genetics
              Yata 1111, Mishima, Shizuoka 411, Japan
              Tel: 81-559-81-6854
              Fax: 81-559-81-6855
              Email: ykohara@lab.nig.ac.jp.

FEATURES     Location/Qualifiers
            1..399
            /organism="Caenorhabditis elegans"
            /strain="CB1489 him-8(e1489)"
            /db_xref="taxon:6239"
            /clone="yk4c4"
            /clone_lib="Yuji Kohara unpublished cDNA"
            /sex="hermaphrodite, male"
            /tissue-type="whole animal"
            /dev_stage="varied"

BASE COUNT  124 a 93 c 95 g 86 t 1 others
ORIGIN

alignment_scores:
  Quality: 52.00      Length: 17
  Ratio: 3.714       Gaps: 0
  Percent Similarity: 82.353 Percent Identity: 70.588

alignment_block:
US-08-653-294-28 x D37592 ..
Align seg 1/1 to: D37592 from: 1 to: 399

2 ArgLeuAlaThrArgLeuAsnGluArgGluAsnLeuArgIleAlaLe 18
|||||
74 AGATCCCGGTAGGTTCACGGAGAGAGATCCCACTGAGGATTATCT 123

18 u 18
|
124 G 124

seq_name: gb_est8:C03945

seq_documentation_block: 232 bp mRNA EST 30-JUL-1996
LOCUS C03945 Human heart cDNA (Ynakamura) Homo sapiens cDNA clone
DEFINITION 3NHC2454, mRNA sequence.
ACCESSION C03945
VERSION C03945.1 GI:1467196
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
Tanaka.T., Oglwara.A., Uchiyama.I., Takagi.T., Yazaki.Y. and
Nakamura.Y.
REFERENCE 1 (bases 1 to 232)
AUTHORS Construction of a normalized directionally cloned cDNA library from
adult heart and analysis of 3040 clones by partial sequencing
Genomics 35 (1), 231-235 (1996)
JOURNAL 96299782
MEDLINE
COMMENT On Oct 24, 1995 this sequence version replaced gi:1040105.
Contact: Yusuke Nakamura
Institute of Medical Science
University of Tokyo
4-6-1, Shirokanedai, Minato-ku, Tokyo 108, Japan
Tel: 81-3-5449-5372
Fax: 81-3-5449-5433
Email: yusuke@ims.u-tokyo.ac.jp.
Location/Qualifiers
1..232
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="3NHC2454"

JOURNAL      Unpublished (1994)
COMMENT      Contact: Yuji Kohara
              Gene Library Lab
              National Institute of Genetics
              Yata 1111, Mishima, Shizuoka 411, Japan
              Tel: 81-559-81-6854
              Fax: 81-559-81-6855
              Email: ykohara@lab.nig.ac.jp.

FEATURES     Location/Qualifiers
            1..399
            /organism="Caenorhabditis elegans"
            /strain="CB1489 him-8(e1489)"
            /db_xref="taxon:6239"
            /clone="yk4c4"
            /clone_lib="Yuji Kohara unpublished cDNA"
            /sex="hermaphrodite, male"
            /tissue-type="whole animal"
            /dev_stage="varied"

BASE COUNT  124 a 93 c 95 g 86 t 1 others
ORIGIN

alignment_scores:
  Quality: 52.00      Length: 17
  Ratio: 3.714       Gaps: 0
  Percent Similarity: 82.353 Percent Identity: 70.588

alignment_block:
US-08-653-294-28 x D37592 ..
Align seg 1/1 to: D37592 from: 1 to: 399

2 ArgLeuAlaThrArgLeuAsnGluArgGluAsnLeuArgIleAlaLe 18
|||||
74 AGATCCCGGTAGGTTCACGGAGAGAGATCCCACTGAGGATTATCT 123

18 u 18
|
124 G 124

seq_name: gb_est8:C03945

seq_documentation_block: 232 bp mRNA EST 30-JUL-1996
LOCUS C03945 Human heart cDNA (Ynakamura) Homo sapiens cDNA clone
DEFINITION 3NHC2454, mRNA sequence.
ACCESSION C03945
VERSION C03945.1 GI:1467196
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
Tanaka.T., Oglwara.A., Uchiyama.I., Takagi.T., Yazaki.Y. and
Nakamura.Y.
REFERENCE 1 (bases 1 to 232)
AUTHORS Construction of a normalized directionally cloned cDNA library from
adult heart and analysis of 3040 clones by partial sequencing
Genomics 35 (1), 231-235 (1996)
JOURNAL 96299782
MEDLINE
COMMENT On Oct 24, 1995 this sequence version replaced gi:1040105.
Contact: Yusuke Nakamura
Institute of Medical Science
University of Tokyo
4-6-1, Shirokanedai, Minato-ku, Tokyo 108, Japan
Tel: 81-3-5449-5372
Fax: 81-3-5449-5433
Email: yusuke@ims.u-tokyo.ac.jp.
Location/Qualifiers
1..232
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="3NHC2454"

JOURNAL      Unpublished (1994)
COMMENT      Contact: Yuji Kohara
              Gene Library Lab
              National Institute of Genetics
              Yata 1111, Mishima, Shizuoka 411, Japan
              Tel: 81-559-81-6854
              Fax: 81-559-81-6855
              Email: ykohara@lab.nig.ac.jp.

FEATURES     Location/Qualifiers
            1..399
            /organism="Caenorhabditis elegans"
            /strain="CB1489 him-8(e1489)"
            /db_xref="taxon:6239"
            /clone="yk4c4"
            /clone_lib="Yuji Kohara unpublished cDNA"
            /sex="hermaphrodite, male"
            /tissue-type="whole animal"
            /dev_stage="varied"

BASE COUNT  124 a 93 c 95 g 86 t 1 others
ORIGIN

alignment_scores:
  Quality: 52.00      Length: 17
  Ratio: 3.714       Gaps: 0
  Percent Similarity: 82.353 Percent Identity: 70.588

alignment_block:
US-08-653-294-28 x D37592 ..
Align seg 1/1 to: D37592 from: 1 to: 399

2 ArgLeuAlaThrArgLeuAsnGluArgGluAsnLeuArgIleAlaLe 18
|||||
74 AGATCCCGGTAGGTTCACGGAGAGAGATCCCACTGAGGATTATCT 123

18 u 18
|
124 G 124

seq_name: gb_est8:C03945

seq_documentation_block: 232 bp mRNA EST 30-JUL-1996
LOCUS C03945 Human heart cDNA (Ynakamura) Homo sapiens cDNA clone
DEFINITION 3NHC2454, mRNA sequence.
ACCESSION C03945
VERSION C03945.1 GI:1467196
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
Tanaka.T., Oglwara.A., Uchiyama.I., Takagi.T., Yazaki.Y. and
Nakamura.Y.
REFERENCE 1 (bases 1 to 232)
AUTHORS Construction of a normalized directionally cloned cDNA library from
adult heart and analysis of 3040 clones by partial sequencing
Genomics 35 (1), 231-235 (1996)
JOURNAL 96299782
MEDLINE
COMMENT On Oct 24, 1995 this sequence version replaced gi:1040105.
Contact: Yusuke Nakamura
Institute of Medical Science
University of Tokyo
4-6-1, Shirokanedai, Minato-ku, Tokyo 108, Japan
Tel: 81-3-5449-5372
Fax: 81-3-5449-5433
Email: yusuke@ims.u-tokyo.ac.jp.
Location/Qualifiers
1..232
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="3NHC2454"

JOURNAL      Unpublished (1994)
COMMENT      Contact: Yuji Kohara
              Gene Library Lab
              National Institute of Genetics
              Yata 1111, Mishima, Shizuoka 411, Japan
              Tel: 81-559-81-6854
              Fax: 81-559-81-6855
              Email: ykohara@lab.nig.ac.jp.

FEATURES     Location/Qualifiers
            1..399
            /organism="Caenorhabditis elegans"
            /strain="CB1489 him-8(e1489)"
            /db_xref="taxon:6239"
            /clone="yk4c4"
            /clone_lib="Yuji Kohara unpublished cDNA"
            /sex="hermaphrodite, male"
            /tissue-type="whole animal"
            /dev_stage="varied"

BASE COUNT  124 a 93 c 95 g 86 t 1 others
ORIGIN

alignment_scores:
  Quality: 52.00      Length: 17
  Ratio: 3.714       Gaps: 0
  Percent Similarity: 82.353 Percent Identity: 70.588

alignment_block:
US-08-653-294-28 x D37592 ..
Align seg 1/1 to: D37592 from: 1 to: 399

2 ArgLeuAlaThrArgLeuAsnGluArgGluAsnLeuArgIleAlaLe 18
|||||
74 AGATCCCGGTAGGTTCACGGAGAGAGATCCCACTGAGGATTATCT 123

18 u 18
|
124 G 124

seq_name: gb_est8:C03945

seq_documentation_block: 232 bp mRNA EST 30-JUL-1996
LOCUS C03945 Human heart cDNA (Ynakamura) Homo sapi
```

```

JOURNAL      Unpublished (1994)
COMMENT      Contact: Yuji Kohara
              Gene Library Lab
              National Institute of Genetics
              Yata 1111, Mishima, Shizuoka 411, Japan
              Tel: 81-559-81-6854
              Fax: 81-559-81-6855
              Email: ykohara@lab.nig.ac.jp.

FEATURES     Location/Qualifiers
Source       1..399
              /organism="Caenorhabditis elegans"
              /strain="CB1489 him-8(e1489)"
              /db_xref="taxon:6239"
              /clone="yk4c4"
              /clone_lib="Yuji Kohara unpublished cDNA"
              /sex="hermaphrodite, male"
              /tissue-type="whole animal"
              /dev_stage="varied"

BASE COUNT   124 a 93 c 95 g 86 t 1 others
ORIGIN

alignment_scores:
  Quality: 52.00      Length: 17
  Ratio: 3.714       Gaps: 0
  Percent Similarity: 82.353 Percent Identity: 70.588

alignment_block:
US-08-653-294-28 x D37592 ..
Align seg 1/1 to: D37592 from: 1 to: 399

2 ArgLeuAlaThrArgLeuAsnGluArgGluAsnLeuArgIleAlaLe 18
|||||
74 AGATCCCGGTAGGTTCACGGAGAGAGATCCCACTGAGGATTATCT 123

18 u 18
|
124 G 124

seq_name: gb_est8:C03945

seq_documentation_block: 232 bp mRNA EST 30-JUL-1996
LOCUS C03945 Human heart cDNA (Ynakamura) Homo sapiens cDNA clone
DEFINITION 3NHC2454, mRNA sequence.
ACCESSION C03945
VERSION C03945.1 GI:1467196
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
Tanaka.T., Oglwara.A., Uchiyama.I., Takagi.T., Yazaki.Y. and
Nakamura.Y.
REFERENCE 1 (bases 1 to 232)
AUTHORS Construction of a normalized directionally cloned cDNA library from
adult heart and analysis of 3040 clones by partial sequencing
Genomics 35 (1), 231-235 (1996)
JOURNAL 96299782
MEDLINE
COMMENT On Oct 24, 1995 this sequence version replaced gi:1040105.
Contact: Yusuke Nakamura
Institute of Medical Science
University of Tokyo
4-6-1, Shirokanedai, Minato-ku, Tokyo 108, Japan
Tel: 81-3-5449-5372
Fax: 81-3-5449-5433
Email: yusuke@ims.u-tokyo.ac.jp.
Location/Qualifiers
Source       1..232
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /clone="3NHC2454"

/clone_lib="Human heart cDNA (Ynakamura)"
/dev_stage="adult"
/note="Organ: heart; normalized directionally cloned cDNA
from adult heart"

BASE COUNT   55 a 77 c 68 g 32 t
ORIGIN

alignment_scores:
  Quality: 50.00      Length: 16
  Ratio: 3.846       Gaps: 0
  Percent Similarity: 81.250 Percent Identity: 68.750

alignment_block:
US-08-653-294-28 x C03945 ..
Align seg 1/1 to: C03945 from: 1 to: 232

5 ThrArgLeuAsnGluArgGluAsnLeuArgIleAlaLeuArgTyr 20
|||||
22 ACCAACACACAGACTTACCGAGAGACCTCGGATCGCGTCGCTAC 69

seq_name: gb_est11:AA263158

seq_documentation_block: 283 bp mRNA EST 02-JUL-1998
LOCUS AA263158 PMY0534 KGI-a Lambda zap Express cDNA library Homo sapiens cDNA 5',
DEFINITION mRNA sequence.
ACCESSION AA263158
VERSION AA263158.1 GI:1898964
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
Claudio.J.O., Liew.C.C., Dempsey.A.A., Cukerman.E., Stewart.A.K.,
Na.E., Atkins.H.I., Iscove.N.N. and Hawley.R.G.
Identification of sequence-tagged transcripts differentially
expressed within the human hematopoietic hierarchy
Genomics 50 (1), 44-52 (1998)
JOURNAL 98292493
MEDLINE
COMMENT On May 5, 1995 this sequence version replaced gi:797810.
Contact: Hawley RG
Oncology Research Laboratories
The Toronto Hospital
CRCS-424, 67 College St., Toronto, Ontario M5G 2M1, Canada
Tel: 416 3403834
Fax: 416 3403453
Email: r.hawley@utoronto.ca
Similar to M8636 MHC class I HLA-Bw gene. Clone was randomly
picked from KGIa primary library
Seq primer: 5' GAATTAACCTCTACTAAAGG 3'
High quality sequence stop: 283.
Location/Qualifiers
Source       1..283
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /clone_lib="KGI-a Lambda Zap Express cDNA library"
              /cell_type="promyeloblast"
              /note="Vector: Lambda zap Express (Stratagene); Site:1:
              EcoRI; Site 2: XhoI; Unidirectional cloning sites:
              EcoRI-XhoI. mRNA was purified from KGI-a cell line, cDNA
              was synthesized using an XhoI-OligodT linker primer. EcoRI
              adaptors were ligated, followed by digestion with XhoI for
              directional cloning into predigested Lambda zap Express"

BASE COUNT   64 a 91 c 88 g 40 t
ORIGIN

alignment_scores:
  Quality: 50.00      Length: 16
  Ratio: 3.846       Gaps: 0
  Percent Similarity: 81.250 Percent Identity: 68.750

alignment_block:
US-08-653-294-28 x C03945 ..
Align seg 1/1 to: C03945 from: 1 to: 232

5 ThrArgLeuAsnGluArgGluAsnLeuArgIleAlaLeuArgTyr 20
|||||
22 ACCAACACACAGACTTACCGAGAGACCTCGGATCGCGTCGCTAC 69

seq_name: gb_est11:AA263158

seq_documentation_block: 283 bp mRNA EST 02-JUL-1998
LOCUS AA263158 PMY0534 KGI-a Lambda zap Express cDNA library Homo sapiens cDNA 5',
DEFINITION mRNA sequence.
ACCESSION AA263158
VERSION AA263158.1 GI:1898964
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
Claudio.J.O., Liew.C.C., Dempsey.A.A., Cukerman.E., Stewart.A.K.,
Na.E., Atkins.H.I., Iscove.N.N. and Hawley.R.G.
Identification of sequence-tagged transcripts differentially
expressed within the human hematopoietic hierarchy
Genomics 50 (1), 44-52 (1998)
JOURNAL 98292493
MEDLINE
COMMENT On May 5, 1995 this sequence version replaced gi:797810.
Contact: Hawley RG
Oncology Research Laboratories
The Toronto Hospital
CRCS-424, 67 College St., Toronto, Ontario M5G 2M1, Canada
Tel: 416 3403834
Fax: 416 3403453
Email: r.hawley@utoronto.ca
Similar to M8636 MHC class I HLA-Bw gene. Clone was randomly
picked from KGIa primary library
Seq primer: 5' GAATTAACCTCTACTAAAGG 3'
High quality sequence stop: 283.
Location/Qualifiers
Source       1..283
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /clone_lib="KGI-a Lambda Zap Express cDNA library"
              /cell_type="promyeloblast"
              /note="Vector: Lambda zap Express (Stratagene); Site:1:
              EcoRI; Site 2: XhoI; Unidirectional cloning sites:
              EcoRI-XhoI. mRNA was purified from KGI-a cell line, cDNA
              was synthesized using an XhoI-OligodT linker primer. EcoRI
              adaptors were ligated, followed by digestion with XhoI for
              directional cloning into predigested Lambda zap Express"

BASE COUNT   64 a 91 c 88 g 40 t
ORIGIN

alignment_scores:
  Quality: 50.00      Length: 16
  Ratio: 3.846       Gaps: 0
  Percent Similarity: 81.250 Percent Identity: 68.750

alignment_block:
US-08-653-294-28 x C03945 ..
Align seg 1/1 to: C03945 from: 1 to: 232

5 ThrArgLeuAsnGluArgGluAsnLeuArgIleAlaLeuArgTyr 20
|||||
22 ACCAACACACAGACTTACCGAGAGACCTCGGATCGCGTCGCTAC 69

seq_name: gb_est11:AA263158

seq_documentation_block: 283 bp mRNA EST 02-JUL-1998
LOCUS AA263158 PMY0534 KGI-a Lambda zap Express cDNA library Homo sapiens cDNA 5',
DEFINITION mRNA sequence.
ACCESSION AA263158
VERSION AA263158.1 GI:1898964
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
Claudio.J.O., Liew.C.C., Dempsey.A.A., Cukerman.E., Stewart.A.K.,
Na.E., Atkins.H.I., Iscove.N.N. and Hawley.R.G.
Identification of sequence-tagged transcripts differentially
expressed within the human hematopoietic hierarchy
Genomics 50 (1), 44-52 (1998)
JOURNAL 98292493
MEDLINE
COMMENT On May 5, 1995 this sequence version replaced gi:797810.
Contact: Hawley RG
Oncology Research Laboratories
The Toronto Hospital
CRCS-424, 67 College St., Toronto, Ontario M5G 2M1, Canada
Tel: 416 3403834
Fax: 416 3403453
Email: r.hawley@utoronto.ca
Similar to M8636 MHC class I HLA-Bw gene. Clone was randomly
picked from KGIa primary library
Seq primer: 5' GAATTAACCTCTACTAAAGG 3'
High quality sequence stop: 283.
Location/Qualifiers
Source       1..283
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /clone_lib="KGI-a Lambda Zap Express cDNA library"
              /cell_type="promyeloblast"
              /note="Vector: Lambda zap Express (Stratagene); Site:1:
              EcoRI; Site 2: XhoI; Unidirectional cloning sites:
              EcoRI-XhoI. mRNA was purified from KGI-a cell line, cDNA
              was synthesized using an XhoI-OligodT linker primer. EcoRI
              adaptors were ligated, followed by digestion with XhoI for
              directional cloning into predigested Lambda zap Express"

BASE COUNT   64 a 91 c 88 g 40 t
ORIGIN

alignment_scores:
  Quality: 50.00      Length: 16
  Ratio: 3.846       Gaps: 0
  Percent Similarity: 81.250 Percent Identity: 68.750

alignment_block:
US-08-653-294-28 x C03945 ..
Align seg 1/1 to: C03945 from: 1 to: 232

5 ThrArgLeuAsnGluArgGluAsnLeuArgIleAlaLeuArgTyr 20
|||||
22 ACCAACACACAGACTTACCGAGAGACCTCGGATCGCGTCGCTAC 69

seq_name: gb_est11:AA263158

seq_documentation_block: 283 bp mRNA EST 02-JUL-1998
LOCUS AA263158 PMY0534 KGI-a Lambda zap Express cDNA library Homo sapiens cDNA 5',
DEFINITION mRNA sequence.
ACCESSION AA263158
VERSION AA263158.1 GI:1898964
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
Claudio.J.O., Liew.C.C., Dempsey.A.A., Cukerman.E., Stewart.A.K.,
Na.E., Atkins.H.I., Iscove.N.N. and Hawley.R.G.
Identification of sequence-tagged transcripts differentially
expressed within the human hematopoietic hierarchy
Genomics 50 (1), 44-52 (1998)
JOURNAL 98292493
MEDLINE
COMMENT On May 5, 1995 this sequence version replaced gi:797810.
Contact: Hawley RG
Oncology Research Laboratories
The Toronto Hospital
CRCS-424, 67 College St., Toronto, Ontario M5G 2M1, Canada
Tel: 416 3403834
Fax: 416 3403453
Email: r.hawley@utoronto.ca
Similar to M8636 MHC class I HLA-Bw gene. Clone was randomly
picked from KGIa primary library
Seq primer: 5' GAATTAACCTCTACTAAAGG 3'
High quality sequence stop: 283.
Location/Qualifiers
Source       1..283
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /clone_lib="KGI-a Lambda Zap Express cDNA library"
              /cell_type="promyeloblast"
              /note="Vector
```

```

Ratio: 3.846      Gaps: 0
Percent Similarity: 81.250      Percent Identity: 68.750

alignment_block:
US-08-653-294-28 x AA263158 ..

Align seg 1/1 to: AA263158 from: 1 to: 283

5 ThrArgLeuAsnGluArgGluAsnLeuArgIleAlaLeuArgTyr 20
|||||:::|||||
102 ACCACACACAGACTTACCGAGAGAACCTGCCGATCCGCTCGGTAC 149

seq_name: gb_est6:D82221

```

seq_documentation_block:					
LOCUS	D82221	375 bp	mRNA	EST	09-FEB-1996
DEFINITION	HUMHC4626 Human pancreatic islet Homo sapiens cDNA similar to HLA-B, mRNA sequence.				
ACCESSION	D82221				
VERSION	D82221.1	GI:1183739			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;				
	Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	1 (bases 1 to 375)				
AUTHORS	Takeda,J.				
TITLE	Human pancreatic islet ESTs				
JOURNAL	Unpublished (1995)				
COMMENT	On Apr 14, 1993 this sequence version replaced gi:785255.				

Contact: Jun Itakeda  
Institute for Molecular and Cellular Regulation, Gunma University  
3-39-15 Showa-machi, Maebashi Gunma 371, Japan  
Tel: 272-20-8856  
Fax: 272-20-8896  
Email: jtakeda@sb.gunma-u.ac.jp.

Location/Qualifiers  
1. .375  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="Human pancreatic islet"  
/note="Vector: Lambda ZAPII; Site\_1: Eco RI; Site\_2: Xho  
I; mRNA was prepared from normal adult human islets. cDNA  
was directionally synthesized from the Xho I in the vector  
to the EcoRI site. cDNA was size fractionated to remove  
sequences <1000 bp in size."

BASE COUNT 75 a 124 c 118 g 55 t 3 others  
ORIGIN

```

alignment_scores:
  Quality: 50.00      Length: 16
  Ratio: 3.846       Gaps: 0
  Percent Similarity: 81.250  Percent Identity: 68.750

alignment_block:
  US-08-653-294-28 x D82221 ..
  Align seq 1/1 to: D82221 from: 1 to: 375

```

```

5 ThrArgLeuAsnGluArgGluAsnLeuArgIleAlaLeuArgTyr 20
||||: :: ||||| ||||| ||||| ||||| |||||
288 ACCAACACACAGACTTACCGAGAGACCTGCGGATCGCGCTCGGTAC 335

seq_name: gb_est31:AI695864

seq_documentation_block:
LOCUS AI695864 748 bp mRNA EST
DEFINITION wc74h11.x1 NCI CGAP_Pan1 Homo sapiens cDNA clone
similar to gb:M28205 H1A CLASS I HISTOCOMPATIBILITY
B-51(B-5) B*5101 ALPHA (HUMAN);, mRNA sequence.
ACCESSION AI695864
VERSION AI695864.1 GI:4984764

```

EST. human.  
SOURCE Homo sapiens  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 748)  
AUTHORS NC1-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT On Mar 16, 1998 this sequence version replaced qi:2961758.

Email: Robert\_Strausberg@nih.gov  
 Life Technologies catalog #: 11548-013  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CCAP clone distribution  
 found through the I.M.A.G.E. Consortium/LLNL at:  
[www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)  
 Contact: Robert Strausberg, Ph.D.  
 Tel: (301) 496-1550

Seq primer: -40UP from Gibco  
High quality sequence stop: 424.

FEATURES	SOURCE
1. <b>Feature 1</b>	Source 1
2. <b>Feature 2</b>	Source 2
3. <b>Feature 3</b>	Source 3
4. <b>Feature 4</b>	Source 4
5. <b>Feature 5</b>	Source 5
6. <b>Feature 6</b>	Source 6
7. <b>Feature 7</b>	Source 7
8. <b>Feature 8</b>	Source 8
9. <b>Feature 9</b>	Source 9
10. <b>Feature 10</b>	Source 10
11. <b>Feature 11</b>	Source 11
12. <b>Feature 12</b>	Source 12
13. <b>Feature 13</b>	Source 13
14. <b>Feature 14</b>	Source 14
15. <b>Feature 15</b>	Source 15
16. <b>Feature 16</b>	Source 16
17. <b>Feature 17</b>	Source 17
18. <b>Feature 18</b>	Source 18
19. <b>Feature 19</b>	Source 19
20. <b>Feature 20</b>	Source 20
21. <b>Feature 21</b>	Source 21
22. <b>Feature 22</b>	Source 22
23. <b>Feature 23</b>	Source 23
24. <b>Feature 24</b>	Source 24
25. <b>Feature 25</b>	Source 25
26. <b>Feature 26</b>	Source 26
27. <b>Feature 27</b>	Source 27
28. <b>Feature 28</b>	Source 28
29. <b>Feature 29</b>	Source 29
30. <b>Feature 30</b>	Source 30
31. <b>Feature 31</b>	Source 31
32. <b>Feature 32</b>	Source 32
33. <b>Feature 33</b>	Source 33
34. <b>Feature 34</b>	Source 34
35. <b>Feature 35</b>	Source 35
36. <b>Feature 36</b>	Source 36
37. <b>Feature 37</b>	Source 37
38. <b>Feature 38</b>	Source 38
39. <b>Feature 39</b>	Source 39
40. <b>Feature 40</b>	Source 40
41. <b>Feature 41</b>	Source 41
42. <b>Feature 42</b>	Source 42
43. <b>Feature 43</b>	Source 43
44. <b>Feature 44</b>	Source 44
45. <b>Feature 45</b>	Source 45
46. <b>Feature 46</b>	Source 46
47. <b>Feature 47</b>	Source 47
48. <b>Feature 48</b>	Source 48
49. <b>Feature 49</b>	Source 49
50. <b>Feature 50</b>	Source 50
51. <b>Feature 51</b>	Source 51
52. <b>Feature 52</b>	Source 52
53. <b>Feature 53</b>	Source 53
54. <b>Feature 54</b>	Source 54
55. <b>Feature 55</b>	Source 55
56. <b>Feature 56</b>	Source 56
57. <b>Feature 57</b>	Source 57
58. <b>Feature 58</b>	Source 58
59. <b>Feature 59</b>	Source 59
60. <b>Feature 60</b>	Source 60
61. <b>Feature 61</b>	Source 61
62. <b>Feature 62</b>	Source 62
63. <b>Feature 63</b>	Source 63
64. <b>Feature 64</b>	Source 64
65. <b>Feature 65</b>	Source 65
66. <b>Feature 66</b>	Source 66
67. <b>Feature 67</b>	Source 67
68. <b>Feature 68</b>	Source 68
69. <b>Feature 69</b>	Source 69
70. <b>Feature 70</b>	Source 70
71. <b>Feature 71</b>	Source 71
72. <b>Feature 72</b>	Source 72
73. <b>Feature 73</b>	Source 73
74. <b>Feature 74</b>	Source 74
75. <b>Feature 75</b>	Source 75
76. <b>Feature 76</b>	Source 76
77. <b>Feature 77</b>	Source 77
78. <b>Feature 78</b>	Source 78
79. <b>Feature 79</b>	Source 79
80. <b>Feature 80</b>	Source 80
81. <b>Feature 81</b>	Source 81
82. <b>Feature 82</b>	Source 82
83. <b>Feature 83</b>	Source 83
84. <b>Feature 84</b>	Source 84
85. <b>Feature 85</b>	Source 85
86. <b>Feature 86</b>	Source 86
87. <b>Feature 87</b>	Source 87
88. <b>Feature 88</b>	Source 88
89. <b>Feature 89</b>	Source 89
90. <b>Feature 90</b>	Source 90
91. <b>Feature 91</b>	Source 91
92. <b>Feature 92</b>	Source 92
93. <b>Feature 93</b>	Source 93
94. <b>Feature 94</b>	Source 94
95. <b>Feature 95</b>	Source 95
96. <b>Feature 96</b>	Source 96
97. <b>Feature 97</b>	Source 97
98. <b>Feature 98</b>	Source 98
99. <b>Feature 99</b>	Source 99
100. <b>Feature 100</b>	Source 100

Source	BASE COUNT	169 a	227 c	237 g	108 t	7 others
/organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:2324421" /clone_lib="NCI CGAP Pan1" /tissue.type="adenocarcinoma" /lab_host="DH108" /note="NCI: pancreas; Vector: pCMV-SPOR16; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.72 kb. Life Technologies catalog #: 11548-013"						

```

alignment_scores:
  Quality: 50.00      Length: 16
             Ratio: 3.846      Gaps: 0
  Percent Similarity: 81.250      Percent Identity: 68.750

alignment_block:
  US-08-653-294-28 x AR696864      ..
  Align seq 1/1 to: AR696864      from: 1 to: 748

```

Align seq 1/1 to: AI696864 from: 1 to: 748

5 ThrArgLeuAsnGluArgArgGluAsnLeuArgIleAlaLeuArgTyr 20  
|||::: |||  
119 ACCAACACACAGACTTACCGAGAGAACCTCGGGATCGCGCTCGCTAC 166

seq name: qb est10:AA151891

seq_documentation_block:	255 bp	mrna	EST	10-DEC-1996
LOCUS	AA151891			
DEFINITION	ZOO1f06_r1 Stragatene colon (#372704)	Hom sapiens	cdna clone	
IMAGE:564435 5'	similar to gb:M15497_cds1 H1A CLASS I			
HISTOCOMPATIBILITY ANTIGEN, A-24(A-9)	A*2401 (HUMAN);			
	sequence.			

sequence.	
ACCESSION	AA151891
VERSION	AA151891.1
	GI:1720754

VERSION  
KEYWORDS  
SOURCE  
EST.  
human

SOURCE ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

1 (bases 1 to 255)

REFERENCE

AUTHORS

Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiappelli, B., Chisnoe, S., Dietrich, N., Dubouquet, T., Favello, A., Gish, W., Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N., Maridis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, R., Schellenger, K., Soares, M.B., Tan, F., Thierry-Mieg, D., Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, D.,

Trevaskis, E., Underwood, K., Wohldmann, P., Waterston, R., Wilson, R. and Marra, M.

Generation and analysis of 280,000 human expressed sequence tags

Genome Res. 6 (9), 807-828 (1996)

97044478

On May 8, 1995 this sequence version replaced gi:800234.

Contact: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Trace considered overall poor quality

Seq primer: -28W13 rev2 from Amersham

High quality sequence stop: 1.

Location/Qualifiers

1. .255

/organism="Homo sapiens"

/db\_xref="CD8:4590888"

/db\_xref="taxon:9606"

/clone="IMAGE:566435"

/clone\_lib="Stratagene colon (#937204)"

/lab\_host="SOLR cells (kanamycin resistant)"

/note="Organ: colon; vector: pBluescript SK-; Site: 1:

ECORI; Site: 2: XhoI; Cloned unidirectionally. Primer:

Oligo dr. T-84 colonic epithelial cell line. Average

insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor

sequence: 5' GAATTCGCGCAGAG 3' -3' adaptor sequence: 5'

CTCGAGTGTGTTTTTTTTTTT 3'"

BASE COUNT 57 a 70 c 75 g 44 t 9 others

ORIGIN

alignment\_scores:

Quality: 49.00 Length: 10

Ratio: 4.900 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-08-653-294-28 x AAL51891

Align seg 1/1 to: AAL51891 from: 1 to: 255

11 ArgGluAsnLeuArgIleAlaLeuArgTyr 20

|||||

77 CGAGAGAACCTCGGATCGCNCCTCGGTAC 106

seq\_name: gb\_est36:AV186046

seq\_documentation\_block:

LOCUS AV186046 360 bp mRNA EST 22-JUL-1999

DEFINITION AV186046 Yuji Kohara unpublished cDNA:Strain N2 hermaphrodite

embryo Caenorhabditis elegans cDNA clone yk494e6 5', mRNA sequence.

AV186046

ACCESSION AV186046.1 GI:5568029

KEYWORDS EST.

SOURCE

ORGANISM

Caenorhabditis elegans.

Caenorhabditis elegans

Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;

Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.

1 (bases 1 to 360)

Kohara, Y., Shin-i, T., Thierry-Mieg, J., Thierry-Mieg, D., Mitsuki, H.,

Nishigaki, A., Motohashi, T., Zeng, Q., Watanabe, H., Sugimoto, A.,

Sano, M., Miyata, A., Mitani, Y., Iida, K., Uesugi, H., Sugiyama, Y. and

Nomoto, H.

Expressed genes in C. elegans

Unpublished (1999)

On Dec 20, 1995 this sequence version replaced gi:1135120.

Contact: Yuji Kohara

Gene Library Lab

National Institute of Genetics

TITLE

JOURNAL

COMMENT

Source

Location/Qualifiers

1. .393

/organism="Caenorhabditis elegans"

/db\_xref="taxon:6239"

/clone="yk494e6"

/clone\_lib="Yuji Kohara unpublished cDNA:Strain N2

hermaphrodite embryo"

/sex="hermaphrodite"

/dev\_stage="embryo"

Yata 1111, Mishima, Shizuoka 411, Japan

Tel: 81-559-81-6854

Fax: 81-559-81-6855

Email: ykohara@lab.nig.ac.jp.

Location/Qualifiers

1. .360

/organism="Caenorhabditis elegans"

/strain="N2"

/db\_xref="taxon:6239"

/clone="yk494e6"

/clone\_lib="Yuji Kohara unpublished cDNA:Strain N2

hermaphrodite embryo"

/sex="hermaphrodite"

/dev\_stage="embryo"

BASE COUNT 103 a 90 c 109 g 58 t

ORIGIN

alignment\_scores:

Quality: 49.00 Length: 19

Ratio: 3.062 Gaps: 0

Percent Similarity: 84.211 Percent Identity: 52.632

alignment\_block:

US-08-653-294-28 x AV186046/rev

Align seg 1/1 to reverse of: AV186046 from: 1 to: 360

2 ArgLeuAlaThrArgLeuAsnGluArgGluAsnLeuArgIleAlaLe 18

174 CGTTGTGCTCAGCGCTCGGGAGCGACGACGATCCGGCTTCGTCT 125

18 uArgTyr 20

|||||

124 CCGATT 118

seq\_name: gb\_est33:AI777709

seq\_documentation\_block:

LOCUS AI777709 393 bp mRNA EST 29-JUN-1999

DEFINITION EST258504 tomato susceptible, Cornell Lycopersicon esculentum cDNA

clone CLES2H21, mRNA sequence.

ACCESSION AI777709

VERSION AI777709.1 GI:5275666

KEYWORDS EST.

SOURCE

ORGANISM

Lycopersicon esculentum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core

eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum;

Potatoe; Lycopersicon.

1 (bases 1 to 393)

D' Ascenzo, M., He, X., Lyman, J., Matern, A.L., Vision, T., Holt, I.E.,

Liang, F., Upton, J., Ronning, C.M., Craven, M.B., Fujii, C.Y.,

Bowman, C.L., Nierman, W., Fraser, C.M., Venter, J.C., Tanksley, S.D.,

Giovannoni, J.J. and Martin, G.B.

Generation of ESTs from Pseudomonas susceptible tomato

Unpublished (1999)

On Jun 5, 1998 this sequence version replaced gi:3187603.

Other ESTs: TC2175

Contact: David Frisch

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Tel: 864 656 4366

Fax: 864 656 4293

Email: dfrisch@clemson.edu

5 prime sequence.

Location/Qualifiers

1. .393

/organism="Lycopersicon esculentum"

/cultivar="Fl-13 (Rio Grande x Money Maker)"

/db\_xref="taxon:4081"

```

/clone="cLES2H21"
/clone_lib="tomato susceptible, Cornell"
/tissue_type="leaf"
/dev_stage="4-week old"
/lab_host="SOLR"
/note="Vector: pBlueScript SK(-); Site_1: EcoRI; Site_2:
XhoI; cLES - Tomato Pseudomonas Susceptible EST Library.
Directionally cloned cDNAs inserted into pBlueScript
SK(-) at 5' end with EcoRI and 3' end with XhoI site"
BASE COUNT      119 a 57 c 70 g 147 t
ORIGIN

alignment_scores:
  Quality: 49.00      Length: 20
  Ratio: 3.267       Gaps: 0
  Percent Similarity: 75.000  Percent Identity: 45.000

alignment_block:
US-08-653-294-28 x AI777709/rev ..
Align seg 1/1 to reverse of: AI777709 from: 1 to: 393

1 TyrArgLeuAlaThrArgLeuAsnGluArgArgGluAsnLeuArgIleAl 17
||||| ||| :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
365 TATATATATTATACAAAGAACAAATACAAACAGAGAAATCTAAATCTAGC 316

17 aLeuArgTyr 20
|:|:|:|:|:|
315 AATCAAAATAT 306

seq_name: gb_est33:AI774937

seq_documentation_block: 496 bp mRNA EST 29-JUN-1999
LOCUS AI774937
DEFINITION EST256037 tomato resistant, Cornell Lycopersicon esculentum cDNA
clone cLER13H12, mRNA sequence.
ACCESSION AI774937
VERSION AI774937.1 GI:5272978
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Potatoe; Lycopersicon.
1 (bases 1 to 496)
D' Ascenzo,M., He,X., Lyman,J., Matern,A.L., Vision,T., Holt,I.E.,
Liang,F., Upton,J., Ronning,C.M., Craven,M.B., Fujii,C.Y.,
Bowman,C.L., Nierman,W., Fraser,C.M., Venter,J.C., Tanksley,S.D.,
Giovannoni,J.J. and Martin,G.B.
Generation of ESTs from Pseudomonas resistant tomato
Unpublished (1999)
On May 18, 1998 this sequence version replaced gi:3136691.
Other ESTs: TC2175
Contact: David Frisch
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 4366
Fax: 864 656 4293
Email: dfrisch@CLEMSON.EDU
5 prime sequence.
FEATURES
  Location/Qualifiers
    1..496
      /organism="Lycopersicon esculentum"
      /cultivar="R11-12 (35S::Pto in Rio Grande x Money Maker)"
      /db_xref="taxon:4081"
      /clone="cLER13H12"
      /clone_lib="tomato resistant, Cornell"
      /tissue_type="leaf"
      /dev_stage="4-week old"
      /lab_host="SOLR"

```

```

/note="Vector: pBlueScript SK(-); Site_1: EcoRI; Site_2:
XhoI; cLER - Tomato Pseudomonas Resistant EST Library.
Directionally cloned cDNAs inserted into pBlueScript
SK(-) at 5' end with EcoRI and 3' end with XhoI site."
BASE COUNT      174 a 88 c 81 g 153 t
ORIGIN

alignment_scores:
  Quality: 49.00      Length: 20
  Ratio: 3.267       Gaps: 0
  Percent Similarity: 75.000  Percent Identity: 45.000

alignment_block:
US-08-653-294-28 x AI774937 ..
Align seg 1/1 to: AI774937 from: 1 to: 496

1 TyrArgLeuAlaThrArgLeuAsnGluArgArgGluAsnLeuArgIleAl 17
||||| ||| :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
18 TATATATATTATACAAAGAACAAATACAAACAGAGAAATCTAAATCTAGC 67

17 aLeuArgTyr 20
|:|:|:|:|:|
68 AATCAAAATAT 77

```

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2000, 04:05:43 ; Search time 133.56 seconds  
(without alignments)  
3.547 Million cell updates/sec

Title: US-08-653-294-29

Perfect score: 99

Sequence: 1 YRLAIRLNERENLTALRY 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 188963 seqs, 23686106 residues

Total number of hits satisfying chosen parameters: 188963

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : A\_Geneseq\_36.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	99	100.0	20	1	R92909	HLA-B2702 CTL modu
2	99	100.0	20	1	W33792	Peptide B2702.84-7
3	94	94.9	20	1	R92910	HLA-B2702 CTL modu
4	94	94.9	20	1	W33793	Peptide B2702.84-7
5	93	93.9	20	1	R92907	HLA-B2702 CTL modu
6	93	93.9	20	1	R95428	HLA-B2702 84-75-84
7	93	93.9	20	1	W33778	Immunomodulating d
8	88	88.9	20	1	R92908	HLA-B2702 CTL modu
9	88	88.9	20	1	W33791	Peptide B2702.84-7
10	80	80.8	20	1	R95430	HLA-B2702 84-75T/7
11	51.5	52.0	20	1	R92911	HLA-B2702 CTL modu
12	51.5	52.0	20	1	W33779	Immunomodulating d
13	50	50.5	10	1	R83095	HLA-B2702 CTL modu
14	50	50.5	10	1	R95426	HLA-B2702.75-84(f)
15	50	50.5	10	1	W33788	Peptide B2702.75-8
16	49	49.5	10	1	W47266	Immunomodulatory p
17	44	44.4	10	1	R41208	Peptide fragment o
18	44	44.4	10	1	R83062	HLA-B2702 CTL modu
19	44	44.4	10	1	R95413	Alpha1-helix of HL
20	44	44.4	10	1	W07512	T-cell modulating
21	44	44.4	10	1	W47265	Immunomodulatory p
22	44	44.4	10	1	W47268	Immunomodulatory p
23	44	44.4	10	1	W47270	Immunomodulatory p
24	44	44.4	10	1	W33784	Peptide B2702.75-8
25	44	44.4	15	1	R92912	HLA-B2702 CTL modu
26	44	44.4	15	1	W33795	Peptide B2702.70-8
27	44	44.4	25	1	R41205	Peptide fragment o
28	44	44.4	25	1	R48286	Peptide fragment o
29	44	44.4	25	1	R83090	HLA-B2702 CTL modu
30	44	44.4	25	1	R83093	HLAB38 CTL modul
31	44	44.4	25	1	R95416	HLA-B2702.60-84. C
32	44	44.4	25	1	R95422	HLAB38.6084. Comps
33	44	44.4	25	1	W33794	Peptide B2702.60-8
34	44	44.4	184	1	Y06801	Peptide seq ID No:

35	44	44.4	362	1	R03142	Sequence of HLA-Bw
36	44	44.4	362	1	R03144	Sequence of HLA-B5
37	44	44.4	362	1	R12463	HLA-Bw53 exon. HLA
38	43	43.4	10	1	W07522	T-cell modulating
39	43	43.4	1849	1	W17900	Photorhabdus lumin
40	43	43.4	1849	1	W56573	Toxin TcdAII, enco
41	43	43.4	2516	1	W17899	Photorhabdus lumin
42	43	43.4	2516	1	W56572	Toxin TcdA, encode
43	42.5	42.9	373	1	Y00067	Enterococcus faeca
44	42.5	42.9	406	1	Y00066	Enterococcus faeca
45	42	42.4	20	1	R92913	HLA-B7 CTL modul

## ALIGNMENTS

### RESULT 1

R92909 R92909 standard; peptide: 20 AA.  
AC R92909;  
DT 16-MAY-1996 (first entry)  
DE HLA-B2702 CTL modulating peptide (B2702.84-75/75-84(T)).  
KW Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;  
KW Immunosuppressant; graft versus host disorder; transplantation; therapy;  
KW class I MHC; HLA-B2702.  
OS Synthetic.  
PN W09526979-Al.  
PD 12-OCT-1995.  
PF 05-APR-1995; 004349.  
PR 05-APR-1994; US-222851.  
PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
PI Clayberger C, Krensky AM, Parham P;  
DR WPI: 95-359582/46.  
PT Extension of acceptance period of transplants from MHC unmatched  
PT donor hosts - using Class I B75-84 MHC antigen of the recipient  
PT host  
PS Example 15; Page 36; 80pp; English.  
CC R83061-R83085, R83090-R83096 and R92907-R92913 represent fragments of  
CC class I major histocompatibility complex (MHC) antigens. This sequence  
CC is an inverted dimer of residues 75-84 of the alpha-1 domain of the class  
CC I MHC HLA-B2702. These sequences can be used to extend the period of  
CC acceptance by a recipient of a transplant from an MHC unmatched donor.  
CC The peptides are administered to a patient in conjunction with a  
CC subtherapeutic amount of an immunosuppressant. This is administered to  
CC the patient for a limited period of time (compared to the lifetime  
CC administration for current treatments). The peptides particularly  
CC modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)  
CC of the patient.  
SQ Sequence 20 AA;

Query Match 100.0%; Score 99; DB 1; Length 20;  
Best Local Similarity 100.0%; Pred. No. 3e-10; Indels 0; Gaps 0;  
Matches 20; Conservative 0; Mismatches 0;

QY 1 YRLAIRLNERENLTALRY 20  
|||||

Db 1 YRLAIRLNERENLTALRY 20  
|||||

### RESULT 2

W33792

ID W33792 standard; peptide: 20 AA.

AC W33792;

DT 19-JUN-1998 (first entry)

DE Peptide B2702.84-75/75-84T tested for immunomodulating activity.

KW Immunomodulating dimer; immunosuppressant drug; CTL activation;

KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;

KW rejection.

OS Synthetic.

OS Homo sapiens.

PN W09744351-Al.

PD 27-NOV-1997.

PF 22-MAY-1997; U08689.  
 PR (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI Beulow R, Clayberger C, Krensky AM;  
 DR WPI; 98-086530/08.  
 PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B  
 PT alpha-1 domain, used for preventing rejection of transplants or  
 PT treating autoimmune diseases  
 PS Example 1; Page 19; 41pp; English.  
 CC Peptides W3784-98 and W3778-9 were assayed for their immunomodulating  
 CC activity. A peptide-type compound or variant is claimed which has  
 CC immunomodulating activity, including the N-terminal acylated and/or  
 CC C-terminal amidated or esterified forms of up to 60 amino acids, where  
 CC the peptide-type compound comprises the formula: A-B, where A, B =  
 CC (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or V; aa77 =  
 CC D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a hydrophobic or  
 CC small amino acid; aa82 = R or L; aa83 = G or R; and aa represents amino  
 CC acid. The sequence in the brackets may optionally be absent or truncated  
 CC at any peptide type bond within the brackets. The compounds comprise  
 CC amino acid sequences related to a Class I HLA-B alpha1 domain (positions  
 CC 79-84). They can be used to inhibit cytotoxic T-lymphocytes (CTL) from  
 CC undesirably attacking cells in a host or in vitro. They can also be  
 CC used in combination with antigenic peptides or proteins of interest to  
 CC activate CTLs. They can also inhibit the proliferation of T cells in  
 CC response to anti-CD3. The peptide can be used for preventing rejection  
 CC of transplants or for treating autoimmune diseases, e.g. diabetes,  
 CC rheumatoid arthritis and lupus erythematosus. The products can also be  
 CC used for detection and diagnosis.  
 SQ Sequence 20 AA;

Query Match 100.0%; Score 99; DB 1; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 3e-10;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YRLAIRLNERENLTALRY 20  
 DB 1 YRLAIRLNERENLTALRY 20  
 ||||| ||||| ||||| ||||| |||||

## RESULT 3

ID R92910 standard; peptide; 20 AA.  
 AC R92910; (first entry)  
 DE HLA-B2702 CTL modulating peptide (B2702.84-75(T)/75-84(T)).  
 DE Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;  
 KW immunosuppressant; graft versus host disorder; transplantation; therapy;  
 KW class I MHC; HLA-B2702.  
 OS Synthetic.  
 PN W09528979-A1.  
 PD 12-OCT-1995.  
 PF 05-APR-1995; U04349.  
 PR (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI Clayberger C, Krensky AM, Parham P;  
 DR WPI; 95-358582/46.  
 PT Extension of acceptance period of transplants from MHC unmatched  
 PT donor hosts - using Class I B75-84 MHC antigen of the recipient  
 PT host  
 PS Example 15; Page 36; 80pp; English.  
 CC R83061-R83085, R83090-R83096 and R92907-R92913 represent fragments of  
 CC class I major histocompatibility complex (MHC) antigens. This sequence  
 CC is an inverted dimer of residues 75-84 of the alpha-1 domain of the class  
 CC I MHC HLA-B2702. These sequences can be used to extend the period of  
 CC acceptance by a recipient of a transplant from an MHC unmatched donor.  
 CC The peptides are administered to a patient in conjunction with a  
 CC subtherapeutic amount of an immunosuppressant. This is administered to  
 CC the patient for a limited period of time (compared to the lifetime  
 CC administration for current treatments). The peptides particularly  
 CC modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)  
 CC of the patient.  
 SQ Sequence 20 AA;

Query Match 94.9%; Score 94; DB 1; Length 20;  
 Best Local Similarity 95.0%; Pred. No. 2e-09;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 YRLAIRLNERENLTALRY 20  
 DB 1 YRLAIRLNERENLTALRY 20  
 ||||| ||||| ||||| ||||| |||||

## RESULT 4

ID W37793 standard; peptide; 20 AA.  
 AC W37793;  
 DT 19-JUN-1998 (first entry)  
 DE Peptide B2702.84-75(T)/75-84(T) tested for immunomodulating activity.  
 KW Immunomodulating dimer; immunosuppressant drug; CTL activation;  
 KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;  
 KW rejection.  
 OS Synthetic.  
 OS Homo sapiens.  
 PN W09744351-A1.  
 PD 27-NOV-1997.  
 PF 22-MAY-1997; U08689.  
 PR 24-MAY-1996; US-653294.  
 PR (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI Beulow R, Clayberger C, Krensky AM;  
 DR WPI; 98-086530/08.  
 PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B  
 PT alpha-1 domain, used for preventing rejection of transplants or  
 PT treating autoimmune diseases  
 PS Example 1; Page 19; 41pp; English.  
 CC Peptides W3784-98 and W3778-9 were assayed for their immunomodulating  
 CC activity. A peptide-type compound or variant is claimed which has  
 CC immunomodulating activity, including the N-terminal acylated and/or  
 CC C-terminal amidated or esterified forms of up to 60 amino acids, where  
 CC the peptide-type compound comprises the formula: A-B, where A, B =  
 CC (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or V; aa77 =  
 CC D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a hydrophobic or  
 CC small amino acid; aa82 = R or L; aa83 = G or R; and aa represents amino  
 CC acid. The sequence in the brackets may optionally be absent or truncated  
 CC at any peptide type bond within the brackets. The compounds comprise  
 CC amino acid sequences related to a Class I HLA-B alpha1 domain (positions  
 CC 79-84). They can be used to inhibit cytotoxic T-lymphocytes (CTL) from  
 CC undesirably attacking cells in a host or in vitro. They can also be  
 CC used in combination with antigenic peptides or proteins of interest to  
 CC activate CTLs. They can also inhibit the proliferation of T cells in  
 CC response to anti-CD3. The peptide can be used for preventing rejection  
 CC of transplants or for treating autoimmune diseases, e.g. diabetes,  
 CC rheumatoid arthritis and lupus erythematosus. The products can also be  
 CC used for detection and diagnosis.  
 SQ Sequence 20 AA;

Query Match 94.9%; Score 94; DB 1; Length 20;  
 Best Local Similarity 95.0%; Pred. No. 2e-09;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 YRLAIRLNERENLTALRY 20  
 DB 1 YRLAIRLNERENLTALRY 20  
 ||||| ||||| ||||| ||||| |||||

## RESULT 5

ID R92907 standard; peptide; 20 AA.  
 AC R92907;  
 DT 16-MAY-1996 (first entry)  
 DE HLA-B2702 CTL modulating peptide (B2702.84-75/75-84).  
 DE Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;  
 KW immunosuppressant; graft versus host disorder; transplantation; therapy;  
 KW class I MHC; HLA-B2702.

OS Synthetic.  
PN W09526979-A1.  
PD 12-OCT-1995.  
PF 05-APR-1995; U04349.  
PR 05-APR-1994; US-222851.  
PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
PI Clayberger C, Krensky AM, Parham P;  
PT WPI: 95-358582/46.  
PT Extension of acceptance period of transplants from MHC unmatched  
PT donor hosts - using Class I B75-84 MHC antigen of the recipient  
PT host  
PS Example 15; Page 36; 80pp; English.  
PS R83085, R83090-R83096 and R92907-R92914 represent fragments of  
CC class I major histocompatibility complex (MHC) antigens. This sequence  
CC is an inverted dimer of residues 75-84 of the alpha-1 domain of the class  
CC I MHC HLA-B\*2702. These sequences can be used to extend the period of  
CC acceptance by a recipient of a transplant from an MHC unmatched donor.  
CC The peptides are administered to a patient in conjunction with a  
CC subtherapeutic amount of an immunosuppressant. This is administered to  
CC the patient for a limited period of time (compared to the lifetime  
CC administration for current treatments). The peptides particularly  
CC modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)  
CC of the patient.  
CC Sequence 20 AA;  
SQ

Query Match 93.9%; Score 93; DB 1; Length 20;  
Best Local Similarity 95.0%; Pred. No. 2.9e-09;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 YRLAIRLNERRENLTALRY 20  
DB 1 YRLAIRLNERRENLTALRY 20  
RESULT 7  
ID W33778  
ID W33778 standard; peptide; 20 AA.  
AC W33778;  
DT 19-JUN-1998 (first entry)  
DE Immunomodulating dimer peptide #1.  
DE Immunomodulating dimer; immunosuppressant drug; CTL activation;  
KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;  
KW rejection.  
OS Synthetic.  
OS Homo sapiens.  
PN W09744351-A1.  
PD 27-NOV-1997.  
PF 22-MAY-1997; U08689.  
PR 24-MAY-1996; US-653294.  
PI (STRD ) UNIV LELAND STANFORD JUNIOR.  
PI Beulow R, Clayberger C, Krensky AM;  
DR WPI: 98-086530/08.  
PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B  
PT alpha-1 domain, used for preventing rejection of transplants or  
PT treating autoimmune diseases  
PS Claim 16; Page 35; 41pp; English.  
PS This sequence represents a specifically claimed immunomodulating  
CC dimer peptide of the invention. A peptide-type compound or variant is  
CC claimed which has immunomodulating activity, including the N-terminal  
CC acylated and/or C-terminal amidated or esterified forms of up to 60  
CC amino acids, where the peptide-type compound comprises the formula: A-B;  
CC where A, B = (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or  
CC V; aa77 = D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a  
CC hydrophobic or small amino acid; aa82 = R or L; aa83 = G or R; and aa  
CC represents amino acid. The sequence in the brackets may optionally be  
CC absent or truncated at any peptide type bond within the brackets. The  
CC compounds comprise amino acid sequences related to a Class I HLA-B  
CC alpha1 domain (positions 79-84). They can be used to inhibit cytotoxic  
CC T-lymphocytes (CTL) from undesirably attacking cells in a host or in  
CC vitro. They can also be used in combination with antigenic peptides or  
CC proteins of interest to activate CTLs. They can also inhibit the  
CC proliferation of T cells in response to anti-CD3. The peptide can be  
CC used for preventing rejection of transplants or for treating autoimmune  
CC diseases, e.g. diabetes, rheumatoid arthritis and lupus erythematosus.  
CC The products can also be used for detection and diagnosis.  
SQ Sequence 20 AA;

Query Match 93.9%; Score 93; DB 1; Length 20;  
Best Local Similarity 95.0%; Pred. No. 2.9e-09;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 YRLAIRLNERRENLTALRY 20  
DB 1 YRLAIRLNERRENLTALRY 20  
RESULT 8  
ID R92908  
ID R92908 standard; peptide; 20 AA.  
AC R92908;  
DT 16-MAY-1996 (first entry)  
DE

OS Synthetic.  
PN W09526979-A1.  
PD 12-OCT-1995.  
PF 05-APR-1995; U04349.  
PR 05-APR-1994; US-222851.  
PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
PI Clayberger C, Krensky AM, Parham P;  
PT WPI: 95-358582/46.  
PT Extension of acceptance period of transplants from MHC unmatched  
PT donor hosts - using Class I B75-84 MHC antigen of the recipient  
PT host  
PS Example 15; Page 36; 80pp; English.  
PS R83085, R83090-R83096 and R92907-R92914 represent fragments of  
CC class I major histocompatibility complex (MHC) antigens. This sequence  
CC is an inverted dimer of residues 75-84 of the alpha-1 domain of the class  
CC I MHC HLA-B\*2702. These sequences can be used to extend the period of  
CC acceptance by a recipient of a transplant from an MHC unmatched donor.  
CC The peptides are administered to a patient in conjunction with a  
CC subtherapeutic amount of an immunosuppressant. This is administered to  
CC the patient for a limited period of time (compared to the lifetime  
CC administration for current treatments). The peptides particularly  
CC modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)  
CC of the patient.  
CC Sequence 20 AA;  
SQ

Query Match 93.9%; Score 93; DB 1; Length 20;  
Best Local Similarity 95.0%; Pred. No. 2.9e-09;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 YRLAIRLNERRENLTALRY 20  
DB 1 YRLAIRLNERRENLTALRY 20  
RESULT 6  
ID R95428  
ID R95428 standard; peptide; 20 AA.  
AC R95428;  
DT 12-NOV-1996 (first entry)  
DE HLA-B\*2702 84-75-84 palindromic.  
DE HLA; p74; alpha-helix; human-leucocyte-associated antigen; inhibitor;  
KW T-cell lysate; membrane protein; mammal; heat shock protein; Hsc70; APC;  
KW B cell; calcium influx; cytotoxic T lymphocyte; CTL; differentiation;  
KW Cytolysis; antigen presenting cell.  
OS Synthetic.  
OS Homo sapiens.  
PN W09513288-A1.  
PD 18-MAY-1995.  
PF 10-NOV-1994; U12985.  
PR 10-NOV-1993; US-150493.  
PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
PI Clayberger C, Krensky AM;  
DR WPI: 95-194027/25.  
PT Compens. comprising lymphoid surface membrane proteins - which may  
PT inhibit cytolytic activity and differentiation of CTLs.  
PS Example; Page 12; 29pp; English.  
PS R95413, and R95415-R95431 represent palindromes and fragments of  
CC human-leucocyte-associated antigens. This sequence represents the  
CC HLA-B\*2702 84-75-84 palindromic. These sequences can be used to isolate  
CC the protein p74 from a T-cell lysate. p74 is a T-cell surface membrane  
CC protein associated with T-cell activation in mammalian T-cells, and is  
CC also immunologically cross reactive with the heat shock protein Hsc70.  
CC p74 is found in a limited number of cell types, but is particularly  
CC expressed on B and T cells. p74 can be isolated by lysis of a suitable  
CC cell with an amphoteric detergent, and then passed through an affinity  
CC column containing a covalently bound HLA-B\*2702 palindromic peptide.  
CC Compositions comprising the extracellular fragment of p74 combined with  
CC HLA-B\*2702.60-84 (see R95416), induces calcium influx, and inhibits  
CC cytotoxic T lymphocyte (CTL) differentiation or cytotoxicity. Candidate  
CC compounds can be screened for their effect on the cytolytic activity of  
CC T-cells, by combining them with the extracellular portion of p74 and  
CC determining the amount of binding between the candidate compound and p74.  
CC Modulation of CTL activity can be inhibited in a cellular composition

DE HLA-B2702 CTL modulating peptide (B2702.84-75(T)/75-84).  
 KW Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;  
 KW immunosuppressant; graft versus host disorder; transplantation; therapy;  
 KW class I MHC; HLA-B2702.  
 OS Synthetic.  
 PN WO9526979-A1.  
 PD 12-OCT-1995.  
 PF 05-APR-1995; U04349.  
 PR 05-APR-1994; US-222851.  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI Clayberger C, Krensky AM, Parham P;  
 DR WPI: 95-358582/46.  
 PT Extension of acceptance period of transplants from MHC unmatched  
 PT donor hosts - using Class I B75-84 MHC antigen of the recipient  
 PT host  
 PS Example 15; Page 36; 80pp; English.  
 CC R83061-R83085, R83090-R83096 and R92907-R92914 represent fragments of  
 CC class I major histocompatibility complex (MHC) antigens. This sequence  
 CC is an inverted dimer of residues 75-84 of the alpha-1 domain of the class  
 CC I MHC HLA-B2702. These sequences can be used to extend the period of  
 CC acceptance by a recipient of a transplant from an MHC unmatched donor.  
 CC The peptides are administered to a patient in conjunction with a  
 CC subtherapeutic amount of an immunosuppressant. This is administered to  
 CC the patient for a limited period of time (compared to the lifetime  
 CC administration for current treatments). The peptides particularly  
 CC modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)  
 CC of the patient.  
 CC Sequence 20 AA;  
 SQ

Query Match 88.9%; Score 88; DB 1; Length 20;  
 Best Local Similarity 90.0%; Pred. No. 2e-08;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 YRLAIRLNERENLRALRY 20  
 ||||| ||||| ||||| |||||  
 Db 1 YRLATRLNERENLRALRY 20

## RESULT 9

ID W33791 standard; peptide; 20 AA.  
 AC W33791;  
 DE Peptide B2702.84-75(T)/75-84 tested for immunomodulating activity.  
 DT Immunomodulating dimer; immunosuppressant drug; CTL activation;  
 KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;  
 KW rejection.  
 OS Homo sapiens.  
 PN WO9744351-A1.  
 PD 27-NOV-1997; U08689.  
 PF 22-MAY-1997; US-653294.  
 PR (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI Beulow R, Clayberger C, Krensky AM;  
 DR WPI: 98-086530/08.  
 PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B  
 PT alpha-1 domain, used for preventing rejection of transplants or  
 PT treating autoimmune diseases  
 PS Example 1; Page 19; 41pp; English.  
 CC Peptides W33784-98 and W33778-9 were assayed for their immunomodulating  
 CC activity. A peptide-type compound or variant is claimed which has  
 CC immunomodulating activity, including the N-terminal acylated and/or  
 CC C-terminal amidated or esterified forms of up to 60 amino acids, where  
 CC the peptide-type compound comprises the formula: A-B, where A, B =  
 CC (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 - E or V; aa77 -  
 CC D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a hydrophobic or  
 CC small amino acid; aa82 = R or L; aa83 = G or R; and aa represents amino  
 CC acid. The sequence in the brackets may optionally be absent or truncated  
 CC at any peptide type bond within the brackets. The compounds comprise  
 CC amino acid sequences related to a Class I HLA-B alpha domain (positions  
 CC 79-84). They can be used to inhibit cytotoxic T-lymphocytes (CTL) from

CC undesirably attacking cells in a host or in vitro. They can also be  
 CC used in combination with antigenic peptides or proteins of interest to  
 CC activate CTLs. They can also inhibit the proliferation of T cells in  
 CC response to anti-CD3. The peptide can be used for preventing rejection  
 CC of transplants or for treating autoimmune diseases, e.g. diabetes,  
 CC rheumatoid arthritis and lupus erythematosus. The products can also be  
 CC used for detection and diagnosis.  
 SQ Sequence 20 AA;

Query Match 88.9%; Score 88; DB 1; Length 20;  
 Best Local Similarity 90.0%; Pred. No. 2e-08;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 YRLAIRLNERENLRALRY 20  
 ||||| ||||| ||||| |||||  
 Db 1 YRLATRLNERENLRALRY 20

## RESULT 10

ID R95430 standard; peptide; 20 AA.  
 AC R95430;  
 DE 12-NOV-1996 (first entry)  
 DT HLA-B2702.84-75(T)/75-84T palindromic.  
 KW HLA; p74; alpha1-helix; human-leucocyte-associated antigen; inhibitor;  
 KW T-cell lysate; membrane protein; mammal; heat shock protein; Hsc70; APC;  
 KW B cell; calcium influx; cytotoxic T lymphocyte; CTL; differentiation;  
 KW cytolysis; antigen presenting cell.  
 OS Synthetic.  
 PN WO9513288-A1.  
 PD 18-MAY-1995.  
 PF 10-NOV-1994; U12985.  
 PR 10-NOV-1993; US-150493.  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI Clayberger C, Krensky AM;  
 DR WPI: 95-194027/25.  
 PT Compens. comprising lymphoid surface membrane proteins - which may  
 PT inhibit cytolytic activity and differentiation of CTLs.  
 PS Example; Page 12; 29pp; English.  
 CC R95413, and R95415-R95431 represent palindromes and fragments of  
 CC human-leucocyte-associated antigens. These sequences represent the  
 CC HLA-B2702.84-75(T)/75-84T palindromic. These sequences can be used to  
 CC isolate the protein p74 from a T-cell lysate. p74 is a T-cell surface  
 CC membrane protein associated with T-cell activation in mammalian T-cells,  
 CC and is also immunologically cross reactive with the heat shock protein  
 CC Hsc70. p74 is found in a limited number of cell types, but is  
 CC particularly expressed on B and T cells. p74 can be isolated by lysis of  
 CC a suitable cell with an amphoteric detergent, and then passed through an  
 CC affinity column containing a covalently bound HLA-B2702 palindromic  
 CC peptide. Compositions comprising the extracellular fragment of p74  
 CC combined with HLA-B2702.60-84 (see R95416), induces calcium influx, and  
 CC inhibits cytotoxic T lymphocyte (CTL) differentiation or cytolysis.  
 CC Candidate compounds can be screened for their effect on the cytolytic  
 CC activity of T-cells, by combining them with the extracellular portion of  
 CC p74 and determining the amount of binding between the candidate compound  
 CC and p74. Modulation of CTL activity can be inhibited in a cellular  
 CC composition containing T-cells and antigen presenting cells (APCs), by  
 CC adding to the mix the extracellular portion of p74, in an amount  
 CC sufficient to compete with p74 for the binding of the p74 ligand.  
 SQ Sequence 20 AA;

Query Match 80.8%; Score 80; DB 1; Length 20;  
 Best Local Similarity 89.5%; Pred. No. 4e-07;  
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 YRLAIRLNERENLRALRY 19  
 ||||| ||||| ||||| |||||  
 Db 1 YRLAIRNETRENLRALRY 19

## RESULT 11

CC represents amino acid. The sequence in the brackets may optionally be absent or truncated at any peptide type bond within the brackets. The compounds comprise amino acid sequences related to a Class I HLA-B alpha domain (positions 79-84). They can be used to inhibit cytotoxic T-lymphocytes (CTL) from undesirably attacking cells in a host or in vitro. They can also be used in combination with antigenic peptides or proteins of interest to activate CTLs. They can also inhibit the proliferation of T cells in response to anti-CD3. The peptide can be used for preventing rejection of transplants or for treating autoimmune diseases, e.g. diabetes, rheumatoid arthritis and lupus erythematosus. The products can also be used for detection and diagnosis.

SQ Sequence 20 AA;

Query Match 52.0%; Score 51.5; DB 1; Length 20;  
Best Local Similarity 68.4%; Pred. No. 0.02;  
Matches 13; Conservative 1; Mismatches 2; Indels 3; Gaps

QY 1 YRLAIRNERENLRPTALR 19  
| | | | | | | | | | | | | |  
Db 1 YRLAIRLNER--YRLAIR 16

RESULT 13

R83095 ID R83095 standard; peptide: 10 AA.  
AC R83095;  
DE 16-MAY-1996 (first entry)  
DT HLA-B2702 CTL modulating peptide (B2702.75-84(T)).  
KW Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;  
KW immunosuppressant; graft versus host disorder; transplantation; therapy;  
KW class I MHC; HLA-B2702.  
OS Synthetic.  
PN WO9526979-A1.  
PD 13-OCT-1995.  
PF 05-APR-1995; U04349.  
PR 03-APR-1994; US-222851.  
PI (STRD ) UNIV LELAND STANFORD JUNIOR.  
PA Clayberger C, Krensky AM, Parham P;  
DR WPL: 95-358582/46.  
PT Extension of acceptance period of transplants from MHC unmatched donor hosts - using Class I B75-84 MHC antigen of the recipient host  
PS Example 14; Page 34; 80pp; English.  
PP R83061-R83085, R83090-R83096 and R92907-R92913 represent fragments of class I major histocompatibility complex (MHC) antigens. This sequence corresponds to residues 75-84 of the alpha-1 domain of the class I MHC HLA-B2702. This sequence showed no inhibitory effect upon cytotoxic T lymphocytes (CTLs). These sequences can be used to extend the period of acceptance by a recipient of a transplant from an MHC unmatched donor. The peptides are administered to a patient in conjunction with a subtherapeutic amount of an immunosuppressant. This is administered to the patient for a limited period of time (compared to the lifetime administration for current treatments). The peptides particularly modulate (or inhibit) the activity of the CTLs of the patient.  
SQ Sequence 10 AA;

Query Match 50.5%; Score 50; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.016;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps

QY 11 RENLRATLYR 20  
| | | | | | | | | |  
Db 1 RENLRATLYR 10

RESULT 14

R95426 ID R95426 standard; peptide: 10 AA.  
AC R95426;  
DT 12-NOV-1996 (first entry)  
DE HLA-B2702.75-84(T).

activity. A peptide-type compound or variant is claimed which has immunomodulating activity, including the N-terminal acylated and/or C-terminal amidated or esterified forms of up to 60 amino acids, where the peptide-type compound comprises the formula; A-B, where A, B = (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or V; aa77 = D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84= a hydrophobic or small amino acid; aa82 = R or L; aa83 = G or R; and aa represents amino acid. The sequence in the brackets may optionally be absent or truncated at any peptide type bond within the brackets. The compounds comprise amino acid sequences related to a Class I HLA-B alpha domain (positions 79-84). They can be used to inhibit cytotoxic T-lymphocytes (CTL) from undesirably attacking cells in a host or in vitro. They can also be used in combination with antigenic peptides or proteins of interest to activate CTLs. They can also inhibit the proliferation of T cells in response to anti-CD3. The peptide can be used for preventing rejection of transplants or for treating autoimmune diseases, e.g. diabetes, rheumatoid arthritis and lupus erythematosus. The products can also be used for detection and diagnosis.

Sequence 10 AA;

Qy	11	RENLR	TALRY	20
		1111111111		
Dd	1	RENLR	TALRY	10

Search completed: February 8, 2000, 04:05:43  
Job time: 9360 sec

Qy	11	RENLR	TALRY	20
		1111111111		
Dd	1	RENLR	TALRY	10

W33788	standard; peptide; 10 AA.
ID	W33788 standard; peptide; 10 AA.
AC	W33788;
DE	19-JUN-1998 (first entry)
DE	Peptide B3702.75-84T80 tested for immunomodulating activity.
DE	Immunomodulating dimer; immunosuppressant drug; CTL activation;
KW	transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;
KW	rejection.
KW	Synthetic.
OS	OS Homo sapiens.
PN	PN W09744351-A1.
PD	27-NOV-1997.
PF	22-MAY-1997; U08689.
PR	24-MAY-1996; US-653294.
PA	(STRD ) UNIV LELAND STANFORD JUNIOR.
PR	Beulow R, Clayberger C, Krensky AM;
DI	WPI: 98-086530/08.
PT	New immunomodulating dimer peptide(s) - based on a Class I HLA-B
PT	alpha-1 domain, used for preventing rejection of transplants or
PT	treating autoimmune diseases
PT	Example 1; Page 19; 41pp; English.
CC	Peptides W33784-98 and W33778-9 were assayed for their immunomodul

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 7, 2000, 18:04:38 ; Search time 111.22 seconds  
(without alignments)  
8.482 Million cell updates/sec

Title: US-08-653-294-29

Perfect score: 99

Sequence: 1 YRLAIRLNERRENLRTALRY 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 142080 seqs, 47169319 residues

Total number of hits satisfying chosen parameters: 142080

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : PIR\_62.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50	50.5	355	2 I37516	HLA-B alpha-chain
2	50	50.5	362	2 S25415	class I histocompa
3	50	50.5	362	2 A45850	MHC class I histoc
4	50	50.5	362	2 I81861	MHC HLA-B*44.2 cha
5	50	50.5	362	2 I34442	MHC class I histoc
6	46	46.5	2077	1 W2BE24	240K tegument prot
7	46	46.5	2078	2 T09326	tegment protein -
8	45	45.5	137	2 I80174	class I histocompa
9	45	45.5	359	1 HLHUB4	MHC class I histoc
10	45	45.5	362	2 I34457	MHC class I lympho
11	44	44.4	273	2 I38509	MHC class I histoc
12	44	44.4	274	2 I54463	MHC HLA-B*38 chain
13	44	44.4	292	2 I57806	MHC H-2K-kml mRNA
14	44	44.4	354	2 I59308	class I histocompa
15	44	44.4	354	2 I80168	class I histocompa
16	44	44.4	354	2 I80167	class I histocompa
17	44	44.4	355	2 I80169	class I histocompa
18	44	44.4	355	2 I80171	class I histocompa
19	44	44.4	359	1 HLHUI2	MHC class I histoc
20	44	44.4	362	1 HLHUB8	MHC class I histoc
21	44	44.4	362	2 B45876	class I histocompa
22	44	44.4	362	2 B30345	MHC class I histoc
23	44	44.4	362	2 JH0541	class I histocompa
24	44	44.4	362	2 JH0539	class I histocompa
25	44	44.4	362	2 JH0540	class I histocompa
26	44	44.4	362	2 A45834	MHC class I histoc
27	44	44.4	362	2 I84486	transmembrane glyc
28	44	44.4	362	2 I62045	gene HLA B-1517 pr
29	44	44.4	362	2 I84490	lymphocyte antigen
30	44	44.4	362	2 I37521	HLA-B*57.2 antigen

#### ALIGNMENTS

##### RESULT 1

I37516

HLA-B alpha-chain - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 04-Oct-1996 #sequence\_revision 04-Oct-1996 #text\_change 23-Jul-1999

C:Accession: I37516

R:Gauchat-Feiss, D.; Breur-Vriesendorp, B.S.; Rufer, N.; Jeannet, M.; Roosnek, E.; Tl

Tissue Antigens 44, 261-264, 1994

A>Title: Sequencing of a novel functional HLA-B\*44 subtype differing in two residues i

A:Reference number: I37516; MUID:95176328

A:Accession: I37516

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-355 <RES>

A:Cross-references: EMBL:X75953; NID:G791007; PIDN:CAA53566.1; PID:G791008

C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 50.5%; Score 50; DB 2; Length 355;

Best Local Similarity 100.0%; Pred. No. 1.4;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 11 RENLRTALRY 20

|||||

Db 99 RENLRTALRY 108

##### RESULT 2

S25415

class I histocompatibility antigen HLA-B\*4403 alpha chain - human

C:Species: Homo sapiens (man)

C>Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 23-Jul-1999

C:Accession: S25415

R:Fleischhauer, K.; Kernan, N.A.; Dupont, B.; Yang, S.Y.

A>Title: The two major subtypes of HLA-B\*44 differ for a single amino acid in codon 15

A:Reference number: S25415; MUID:91335451

A:Accession: S25415

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-362 <FILE>

A:Cross-references: EMBL:X64366; NID:G32178; PIDN:CAA45718.1; PID:G32179

C:Genetics:

A:Gene: GDB:HLA-B

A:Cross-references: GDB:120048; OMIM:142830

A:Map position: 6p21.3-6p21.3

C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

C:Keywords: transmembrane protein

F:220-285/Domain: immunoglobulin homology <IMM>

Query Match

Best Local Similarity 100.0%; Score 50; DB 2; Length 362;

Pred. No. 1.5;

Matches 10: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RNLRTALRY 20  
|||||

Db 99 RNLRTALRY 108

# RESULT 3

A45850

MHC class I histocompatibility antigen HLA-B13.1 - human

C:Species: Homo sapiens (man)

C:Date: 03-Jun-1993 #sequence\_revision 03-Jun-1993 #text\_change 23-Jul-1999

C:Accession: A45850

R:Kato, K.; Dupont, B.; Yang, S.Y.

Immunogenetics 29, 117-120, 1989

A:Title: Localization of nucleotide sequence which determines mongoloid subtype of HLA-B

A:Reference number: A45850; MUID:89122134

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-362 <KAT>

A:Cross-references: GB:M24075; NID:g187703; PIDN:AAA59627.1; PID:g386884

C:Genetics:

A:Gene: GDB:HLA-B

A:Cross-references: GDB:120048; OMIM:142830

A:Map position: 6p21.3-6p21.3

C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

F:220-285/Domain: immunoglobulin homology <IMM>

Query Match 50.5%; Score 50; DB 2; Length 362;

Best Local Similarity 100.0%; Pred. No. 1.5;

Matches 10: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RNLRTALRY 20

|||||

Db 99 RNLRTALRY 108

# RESULT 4

I61861

MHC HLA-B44.2 chain - human

C:Species: Homo sapiens (man)

C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 23-Jul-1999

C:Accession: I61861

R:Parham, P.; Lawlor, D.A.; Lomen, C.E.; Ennis, P.D.

J. Immunol. 142, 3937-3950, 1989

A:Title: Diversity and diversification of HLA-A,B,C alleles.

A:Reference number: I36956; MUID:89235215

A:Accession: I61861

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-362 <RES>

A:Cross-references: GB:M24038; NID:g187811; PIDN:AAA59663.1; PID:g386900

C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 50.5%; Score 50; DB 2; Length 362;

Best Local Similarity 100.0%; Pred. No. 1.5;

Matches 10: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RNLRTALRY 20

|||||

Db 99 RNLRTALRY 108

# RESULT 5

I54442

MHC class I histocompatibility antigen HLA-B13 precursor - human

C:Species: Homo sapiens (man)

C:Date: 07-Jun-1996 #sequence\_revision 07-Jun-1996 #text\_change 23-Jul-1999

C:Accession: I54442; I61858

R:Zemmou, J.; Ennis, P.D.; Parham, P.; Dupont, B.

Immunogenetics 27, 281-287, 1988

A:Title: Comparison of the structure of HLA-Bw47 to HLA-B13 and its relationship to 2

A:Reference number: I54442; MUID:88152906

A:Accession: I54442

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-362 <ZEM>

A:Cross-references: GB:M19757; NID:g184161; PIDN:AAA52657.1; PID:g386774

R:Parham, P.; Lawlor, D.A.; Lomen, C.E.; Ennis, P.D.

J. Immunol. 142, 3937-3950, 1989

A:Title: Diversity and diversification of HLA-A,B,C alleles.

A:Reference number: I36956; MUID:89235215

A:Accession: I61858

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-362 <PAR>

A:Cross-references: GB:M24041; NID:g187805; PIDN:AAA59660.1; PID:g386897

C:Genetics:

A:Gene: GDB:HLA-B

A:Cross-references: GDB:120048; OMIM:142830

A:Map position: 6p21.3-6p21.3

C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 50.5%; Score 50; DB 2; Length 362;

Best Local Similarity 100.0%; Pred. No. 1.5;

Matches 10: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RNLRTALRY 20

|||||

Db 99 RNLRTALRY 108

# RESULT 6

WZBEZ4

240K tegument protein (clone pZVH14) - human herpesvirus 6 (strain GS)

C:Species: human herpesvirus 6

A:Note: host Homo sapiens (man)

C:Date: 31-Dec-1992 #sequence\_revision 31-Dec-1992 #text\_change 16-Jul-1999

C:Accession: H40511

R:Josephs, S.F.; Ablashi, D.V.; Salahuddin, S.Z.; Jagodzinski, L.L.; Wong-Staal, F.;

J. Virol. 65, 5597-5604, 1991

A:Title: Identification of the human herpesvirus 6 glycoprotein H and putative large

A:Reference number: A40511; MUID:91374623

A:Molecule type: DNA

A:Residues: 1-2077 <JOS>

A:Cross-references: GB:S57540; NID:g235435; PIDN:AAB19786.1; PID:g235439

C:Superfamily: varicella-zoster virus 240K tegument protein

Query Match 46.5%; Score 46; DB 1; Length 2077;

Best Local Similarity 38.9%; Pred. No. 41;

Matches 7; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

QY 1 YRLAIRLNRERENLRAL 18

|||||

Db 495 YKVAVLLNEKRKEIQEAI 512

# RESULT 7

T09326

tegument protein - human herpesvirus 6 (strain UI102)

C:Species: human herpesvirus 6

C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 23-Jul-1999

C:Accession: T09326

R:Nicholas, J.; Martin, M.

J. Virol. 68, 597-610, 1994

A:Title: Nucleotide sequence analysis of a 38.5-kilobase-pair region of the genome of

A:Reference number: Z16644; MUID:94118404

A:Accession: T09326

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA



A:Residues: 1-2078 <NIC>  
A:Cross-references: EMBL:L25528; NID:q451932; PID:q451957  
C:Genetics:  
A:Gene: HHRF1  
C:Superfamily: varicella-zoster virus 240K tegument protein

Query Match 46.5%; Score 46; DB 2; Length 2078;  
Best Local Similarity 38.9%; Pred. No. 41;  
Matches 7; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

Qy 1 YRLAIRNERRENLTAL 18  
Db 496 YKVAVLLNEKRKEIQEI 513

RESULT 8

MHC class I histocompatibility antigen - chimpanzee (fragment)  
C:Species: Pan troglodytes (chimpanzee)  
C:Date: 24-May-1996 #sequence\_revision 24-May-1996 #text\_change 23-Jul-1999  
C:Accession: I80174  
R:McAdam, S.N.; Boyson, J.E.; Liu, X.; Garber, T.L.; Hughes, A.L.; Bontrop, R.E.; Watkins  
Proc. Natl. Acad. Sci. U.S.A. 91, 5893-5897, 1994  
A:Title: A uniquely high level of recombination at the HLA-B locus.  
A:Reference number: I59308; MUID:94286544  
A:Accession: I80174  
A:Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: DNA  
A:Residues: 1-137 <RES>  
A:Cross-references: EMBL:U05585; NID:q454787; PIDN:AAA50188.1; PID:q454788  
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 45.5%; Score 45; DB 2; Length 137;  
Best Local Similarity 90.0%; Pred. No. 3.2;  
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 11 RNLRLTALRY 20  
Db 40 RNLRLTLLRY 49

RESULT 9

MHC class I histocompatibility antigen HLA-B44 alpha chain precursor - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 31-Dec-1988 #sequence\_revision 31-Dec-1988 #text\_change 22-Jun-1999  
C:Accession: A25295  
R:Kottmann, A.H.; Seemann, G.H.A.; Guessow, H.D.; Roos, M.H.  
Immunogenetics 23, 396-400, 1986  
A:Title: DNA sequence of the coding region of the HLA-B44 gene.  
A:Reference number: A25295; MUID:86249389  
A:Accession: A25295  
A:Molecule type: mRNA  
A:Residues: 1-359 <KOT>  
A:Cross-references: GB:M15470; NID:q187680; PIDN:AAA59619.1; PID:q386883  
C:Genetics:

A:Gene: GDB:HLA-B  
A:Cross-references: GDB:120048; OMIM:142830  
A:Map position: 6p21.3-6p21.3  
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology  
C:Keywords: duplication; glycoprotein; heterodimer; transmembrane protein; transplantati  
F:1-21/Domain: signal sequence (fragment) #status predicted <SIG>  
F:22-359/Product: class I histocompatibility antigen HLA-B44 alpha chain #status predict  
F:22-304/Domain: extracellular alpha-1 <EX1>  
F:22-111/Domain: alpha-1 <EX1>  
F:112-203/Domain: alpha-2 <EX2>  
F:217-282/Domain: immunoglobulin homology <IMM>  
F:305-328/Domain: transmembrane #status predicted <TM>  
F:329-359/Domain: intracellular #status predicted <INT>  
F:107/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 45.5%; Score 45; DB 1; Length 359;  
Best Local Similarity 90.0%; Pred. No. 9;  
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 11 RNLRLTALRY 20  
Db 96 RNLRLTAARY 105

RESULT 10

MHC class I lymphocyte antigen - human  
C:Species: Homo sapiens (man)  
C:Date: 07-Jun-1996 #sequence\_revision 07-Jun-1996 #text\_change 23-Jul-1999  
C:Accession: I54457  
R:Pohla, H.; Kuon, W.; Tabaczewski, P.; Doerner, C.; Weiss, E.H.  
Immunogenetics 29, 297-307, 1989  
A:Title: Allelic variation in HLA-B and HLA-C sequences and the evolution of the HLA-  
A:Reference number: I54457; MUID:89233295  
A:Accession: I54457  
A:Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: mRNA  
A:Residues: 1-362 <RES>  
A:Cross-references: GB:M28205; NID:q576470; PIDN:AAA57145.1; PID:q576471  
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 45.5%; Score 45; DB 2; Length 362;  
Best Local Similarity 90.0%; Pred. No. 9.1;  
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 11 RNLRLTALRY 20  
Db 99 RNLRLTLLRY 108

RESULT 11

MHC class I histocompatibility antigen - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 06-Sep-1996 #sequence\_revision 06-Sep-1996 #text\_change 23-Jul-1999  
C:Accession: I38509  
R:Cereb, N.; Choi, J.W.; Riu, K.Z.; Yang, S.Y.  
Tissue Antigens 44, 271-273, 1994  
A:Title: HLA-B\*5105, a newly identified B51 IEF variant.  
A:Reference number: I38509; MUID:95176331  
A:Accession: I38509  
A:Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: mRNA  
A:Residues: 1-273 <RES>  
A:Cross-references: EMBL:U06697; NID:q469544; PIDN:AAA92997.1; PID:q469545  
C:Genetics:  
A:Gene: GDB:HLA-B  
A:Cross-references: GDB:120048; OMIM:142830  
A:Map position: 6p21.3-6p21.3  
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 44.4%; Score 44; DB 2; Length 273;  
Best Local Similarity 90.0%; Pred. No. 9.7;  
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 11 RNLRLTALRY 20  
Db 74 RNLRLTALRY 83

RESULT 12

MHC HLA-B38 chain - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 07-Jun-1996 #sequence\_revision 07-Jun-1996 #text\_change 23-Jul-1999

C:Accession: I54463  
 R:Mueller, C.A.; Engler-Blum, G.; Gekeler, V.; Steiert, I.; Weiss, E.; Schmidt, H.  
 Immunogenetics 30, 200-207, 1989  
 A:Title: Genetic and serological heterogeneity of the supertypic HLA-B locus specificity  
 A:Reference number: I54463; MUID:89379286  
 A:Accession: I54463  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-274 <RES>

A:Cross-references: GB:M29864; NID:gl87674; PIDN:AAA36222.1; PID:gl87675  
 C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 44.4%; Score 44; DB 2; Length 274;  
 Best Local Similarity 90.0%; Pred. No. 9.7;  
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 11 RENLRLALRY 20  
 ||||| |||||  
 DB 75 RENLRLALRY 84

RESULT 13  
 I57806  
 MHC H-2K-kml mRNA - mouse (fragment)  
 C:Species: Mus musculus (house mouse)  
 C:Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 23-Jul-1999  
 C:Accession: I57806  
 R:Martinko, J.M.; Solheim, J.C.; Gellibter, J.  
 Mol. Immunol. 25, 267-274, 1988  
 A:Title: The H-2K-kml mutation: A single nucleotide substitution is responsible for mult  
 A:Reference number: I57806; MUID:88232726  
 A:Accession: I57806  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-292 <RES>  
 A:Cross-references: GB:M34932; NID:gl99405; PIDN:AAA39596.1; PID:gl99406  
 C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 44.4%; Score 44; DB 2; Length 292;  
 Best Local Similarity 73.3%; Pred. No. 10;  
 Matches 11; Conservative 1; Mismatches 1; Indels 2; Gaps 1;

QY 8 NER--RENLRLALRY 20  
 ||: | ||||| |||||  
 DB 70 NEQIFRVNRLALRY 84

RESULT 14  
 I59308  
 Class I histocompatibility antigen - pygmy chimpanzee (fragment)  
 C:Species: Pan paniscus (pygmy chimpanzee, bonobo)  
 C:Date: 31-May-1996 #sequence\_revision 31-May-1996 #text\_change 23-Jul-1999  
 C:Accession: I59308  
 R:McAdam, S.N.; Boyson, J.E.; Liu, X.; Garber, T.L.; Hughes, A.L.; Bontrop, R.E.; Watkin  
 Proc. Natl. Acad. Sci. U.S.A. 91, 5893-5897, 1994  
 A:Title: A uniquely high level of recombination at the HLA-B locus.  
 A:Reference number: I59308; MUID:94286544  
 A:Accession: I59308  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-354 <RES>  
 A:Cross-references: EMBL:U05575; NID:9454767; PIDN:AAA50178.1; PID:9454768  
 C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 44.4%; Score 44; DB 2; Length 354;  
 Best Local Similarity 90.0%; Pred. No. 13;  
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 11 RENLRLALRY 20  
 ||||| ||||| |||||

DB 91 RENLRLALRY 100 -

RESULT 15

I80168  
 Class I histocompatibility antigen - chimpanzee (fragment)  
 C:Species: Pan troglodytes (chimpanzee)  
 C:Date: 24-May-1996 #sequence\_revision 24-May-1996 #text\_change 23-Jul-1999  
 C:Accession: I80168  
 R:McAdam, S.N.; Boyson, J.E.; Liu, X.; Garber, T.L.; Hughes, A.L.; Bontrop, R.E.; Wat  
 Proc. Natl. Acad. Sci. U.S.A. 91, 5893-5897, 1994  
 A:Title: A uniquely high level of recombination at the HLA-B locus.  
 A:Reference number: I59308; MUID:94286544

A:Accession: I80168  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-354 <RES>  
 A:Cross-references: EMBL:U05579; NID:9454775; PIDN:AAA50182.1; PID:9454776  
 C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 44.4%; Score 44; DB 2; Length 354;  
 Best Local Similarity 90.0%; Pred. No. 13;  
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 11 RENLRLALRY 20  
 ||||| |||||  
 DB 91 RENLRLALRY 100

Search completed: February 7, 2000, 18:04:39  
 Job time: 22205 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2000, 01:25:59 ; Search time 68.63 Seconds  
(without alignments)  
8.703 Million cell updates/sec

Title: US-08-653-294-29

Perfect score: 99

Sequence: 1 YRLAIRLNERENLTALRY 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 82229 seqs, 29864866 residues

Total number of hits satisfying chosen parameters: 82229

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : SwissProt\_38.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50	50.5	362	1 B05 HUMAN	P30461 homo sapien
2	50	50.5	362	1 B41 HUMAN	P30481 homo sapien
3	50	50.5	362	1 B42 HUMAN	P30482 homo sapien
4	46	46.5	2077	1 TEGU HSV6G	P30002 herpes simp
5	46	46.5	2077	1 TEGU HSV6U	P52340 herpes simp
6	45	45.5	359	1 B40 HUMAN	P10320 homo sapien
7	44	44.4	351	1 CBIG SALTY	Q05631 salmonella
8	44	44.4	359	1 B01 PANTR	P13750 pan troglod
9	44	44.4	362	1 B01 GORGO	P30379 gorilla gor
10	44	44.4	362	1 B02 GORGO	P30380 gorilla gor
11	44	44.4	362	1 B03 GORGO	P30381 gorilla gor
12	44	44.4	362	1 B15 HUMAN	P10317 homo sapien
13	44	44.4	362	1 B47 HUMAN	P30487 homo sapien
14	44	44.4	362	1 B49 HUMAN	P18464 homo sapien
15	44	44.4	362	1 B52 HUMAN	P30489 homo sapien
16	44	44.4	362	1 B53 HUMAN	P30490 homo sapien
17	44	44.4	362	1 B54 HUMAN	P30491 homo sapien
18	44	44.4	362	1 B60 HUMAN	P18465 homo sapien
19	44	44.4	362	1 B61 HUMAN	P30497 homo sapien
20	44	44.4	362	1 B62 HUMAN	P10319 homo sapien
21	44	44.4	362	1 HLAH HUMAN	P01893 homo sapien
22	44	44.4	365	1 A23 HUMAN	P30447 homo sapien
23	44	44.4	365	1 A24 HUMAN	P05534 homo sapien
24	44	44.4	369	1 HAIK MOUSE	P04223 mus musculu
25	42	42.4	591	1 GLWS AQUAE	O66848 a glucosam
26	40	40.4	338	1 B20 HUMAN	P30467 homo sapien
27	40	40.4	361	1 B14 HUMAN	P03989 homo sapien
28	40	40.4	362	1 B16 HUMAN	P19373 homo sapien
29	40	40.4	362	1 B18 HUMAN	P10318 homo sapien
30	40	40.4	362	1 B19 HUMAN	Q08136 homo sapien
31	40	40.4	362	1 B29 HUMAN	P18463 homo sapien
32	40	40.4	362	1 B45 HUMAN	P30485 homo sapien
33	40	40.4	433	1 HEM2 SPIOL	P24493 spinacia ol
34	40	40.4	924	1 ORC1 DROME	O16810 drosophila

35 40 40.4 1126 1 MEM2 DROME P55162 drosophila  
36 39 39.4 162 1 PUR6 BACSU P12044 bacillus su  
37 39 39.4 187 1 ATPE ANASP P12407 anabaena sp  
38 39 39.4 281 1 APE BRARE O42364 brachydanio  
39 39 39.4 298 1 HAIY MOUSE P01895 mus musculu  
40 39 39.4 303 1 DMA TREPA O33844 treponema p  
41 39 39.4 365 1 IA25 HUMAN P18462 homo sapien  
42 39 39.4 365 1 IA32 HUMAN P10314 homo sapien  
43 39 39.4 365 1 HAI2 MOUSE P01900 mus musculu  
44 39 39.4 375 1 SR55 DROME P26686 drosophila  
45 39 39.4 380 1 LEU3 PHACH O59930 phanerochaete

#### ALIGNMENTS

RESULT 1  
ID B05\_HUMAN STANDARD; PRT; 362 AA.  
AC P30461;  
DT 01-APR-1993 (Rel. 25, Created)  
DT 01-APR-1993 (Rel. 25, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-13 B\*1301 ALPHA CHAIN  
DE PRECURSOR (B13.1).  
GN HLA-B OR HLAB.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 89235215.  
RA PARHAM P., LAWOR D.A., LOMEN C.E., ENNIS P.D.;  
RT "Diversity and diversification of HLA-A,B,C alleles."  
RL J. Immunol. 142:3937-3950(1989).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 88152906.  
RA ZEMOUR J., ENNIS P.D., PARHAM P., DUPONT B.;  
RT "Comparison of the structure of HLA-B\*47 to HLA-B13 and its  
RL relationship to 21-hydroxylase deficiency."  
RN Immunogenetics 27:281-287(1988).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 96053518.  
RA LIN L., TOKUNAGA K., NAKAJIMA F., ISHIKAWA Y., KASHIWASE K.,  
RA TANAKA H., KUWATA S., SDELTSEVA E., AKAZA T., TADOKORO K.,  
RA SHIBATA Y., CHANDANAVONGYONG D., JUJI T.;  
RT "Both HLA-B\*1301 and B\*1302 exist in Asian populations and are  
RL associated with different haplotypes."  
Hum. Immunol. 43:51-56(1995).  
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO  
CC THE IMMUNE SYSTEM.  
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
CC MICROGLOBULIN).  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; M24041; AAA59660.1; -  
CC DR EMBL; M19757; AAA52657.1; -  
CC DR EMBL; P50291; BAA08822.1; -  
CC DR HSSP; P30491; 1A1M.  
CC MIM; 142830; -  
CC DR PROSITE; PS00290; IG\_MHC; 1.  
CC DR PFAM; PF00047; 1g; 1.  
CC DR PFAM; PF00129; MHC\_I; 1.

KW MHC I; Transmembrane; Glycoprotein; Signal.  
 FT SIGNAL 1 24  
 FT CHAIN 25 362 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,  
 B-13 B\*1301 ALPHA CHAIN.  
 FT DOMAIN 25 114  
 FT DOMAIN 115 206  
 FT DOMAIN 207 298  
 FT DOMAIN 299 309  
 FT DOMAIN 310 333  
 FT TRANSMEM 334 362  
 FT DOMAIN 110 110  
 FT CARBOHYD 125 188  
 FT DISULFID 227 283  
 FT SEQUENCE 362 AA; 40474 MW; 28867875 CRC32;  
 SO

Query Match 50.5%; Score 50; DB 1; Length 362;  
 Best Local Similarity 100.0%; Pred. No. 0.58;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 RENLRTALRY 20  
 |||||  
 Db 99 RENLRTALRY 108

RESULT 2  
 1B41\_HUMAN STANDARD; PRT; 362 AA.  
 ID 1B41\_HUMAN  
 AC P30481;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, BW-44(B-12) B\*4402 ALPHA CHAIN  
 DE PRECURSOR (B44.2).  
 GN HLA-B OR HLAB.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 89235215.  
 RA PARHAM P., LAWLOD D.A., LOMEN C.E., ENNIS P.D.;  
 "Diversity and diversification of HLA-A,B,C alleles.";  
 RL J. Immunol. 142:3937-3950(1989).  
 CC -1- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO  
 THE IMMUNE SYSTEM.  
 CC -1- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
 MICROGLOBULIN).  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; M24038; AAA59663.1;  
 DR HSSP; P30491; IALM.  
 DR MIM; 142830;  
 DR PROSITE; PS00290; IG\_MHC; 1.  
 DR PFAM; PF00047; Ig; 1.  
 DR PFAM; PF00129; MHC\_I; 1.  
 KW MHC I; Transmembrane; Glycoprotein; Signal.  
 FT SIGNAL 1 24  
 FT CHAIN 25 362 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,  
 BW-44(B-12) B\*4402 ALPHA CHAIN.  
 FT DOMAIN 25 114  
 FT DOMAIN 115 206  
 FT DOMAIN 207 298  
 FT DOMAIN 299 309  
 FT DOMAIN 310 333  
 FT TRANSMEM 334 362  
 FT DOMAIN 110 110  
 FT CARBOHYD 125 188  
 FT DISULFID 227 283  
 FT SEQUENCE 362 AA; 40474 MW; 28867875 CRC32;  
 SO

FT CARBOHYD 110 110 BY SIMILARITY.  
 FT DISULFID 125 188 BY SIMILARITY.  
 FT DISULFID 227 283 BY SIMILARITY.  
 SQ SEQUENCE 362 AA; 40481 MW; DOAE6DD5 CRC32;  
 SO

Query Match 50.5%; Score 50; DB 1; Length 362;  
 Best Local Similarity 100.0%; Pred. No. 0.58;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 RENLRTALRY 20  
 |||||  
 Db 99 RENLRTALRY 108

RESULT 3  
 1B42\_HUMAN STANDARD; PRT; 362 AA.  
 ID 1B42\_HUMAN  
 AC P30482;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, BW-44(B-12) B\*4403 ALPHA CHAIN  
 DE PRECURSOR.  
 GN HLA-B OR HLAB.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 91335451.  
 RA FLEISCHHAUER K., KERNAN N.A., DUPONT B., YANG S.Y.;  
 "The two major subtypes of HLA-B44 differ for a single amino acid in  
 codon 156.";  
 RL Tissue Antigens 37:133-137(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 96435470.  
 RA ADAMS E.J., LITTLE A.M., ARNETT K.L., MCAULEY J.E., WILLIAMS R.C.,  
 PARHAM P.;  
 "Three new HLA-B alleles found in Mexican-Americans.";  
 RL Tissue Antigens 46:414-416(1995).  
 CC -1- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO  
 THE IMMUNE SYSTEM.  
 CC -1- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
 MICROGLOBULIN).  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; X64366; CAA45718.1;  
 DR EMBL; L42282; AAB51454.1;  
 DR EMBL; L42283; AAB51455.1;  
 DR PIR; S25415; S25415.  
 DR HSSP; P30491; IALM.  
 DR MIM; 142830;  
 DR PROSITE; PS00290; IG\_MHC; 1.  
 DR PFAM; PF00047; Ig; 1.  
 DR PFAM; PF00129; MHC\_I; 1.  
 KW MHC I; Transmembrane; Glycoprotein; Signal.  
 FT SIGNAL 1 24  
 FT CHAIN 25 362 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,  
 BW-44(B-12) B\*4403 ALPHA CHAIN.  
 FT DOMAIN 25 114  
 FT DOMAIN 115 206  
 FT DOMAIN 207 298  
 FT DOMAIN 299 309  
 FT TRANSMEM 310 333  
 FT DOMAIN 110 110  
 FT CARBOHYD 125 188  
 FT DISULFID 227 283  
 FT SEQUENCE 362 AA; 40481 MW; DOAE6DD5 CRC32;  
 SO

DR PFAM: PF00047; Ig; 1.  
DR PFAM: PF00129; MHC I; 1.

QY 3 LAIRLNERRRNLRTALR 19  
 11 : 11 1 : 1111 :  
 Db 144 LAFOLNARNSDRLRTAVK 160

RESULT 9  
 1B01\_GORG0  
 STANDARD; PRT; 362 AA.  
 ID 1B01\_GORG0  
 AC P30379;

```

DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-APR-1993 (Rel. 25, Last annotation update)
DE CLASS I HISTOCOMPATIBILITY ANTIGEN, GOGO-B0101 ALPHA CHAIN PRECURSOR.
OS Gorilla gorilla gorilla (Lowland gorilla)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
RN [1]
RN SEQUENCE FROM N.A.
RP MEDLINE: 92078860.
RX LAWLOR D.A., WARREN E., TAYLOR P., PARHAM P.:
RA "Corilla class I major histocompatibility complex alleles: comparison
RT to human and chimpanzee class I."
RL J. Exp. Med. 174:1491-1509(1991).
CC -I- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC THE IMMUNE SYSTEM.
CC -I- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC MICROGLOBULIN).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X60255; CAA42807.1; -.
DR PIR: JH0539; JH0539.
DR HSP: P03989; IHS.A.
DR PROSITE: PS00290; IG_MHC; 1.
DR PFAM: PF00047; Ig; 1.
DR PFAM: PF00129; MHC_I; 1.
DR MHC I; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 24
FT CHAIN 25 362
FT FT BY SIMILARITY.
FT FT CLASS I HISTOCOMPATIBILITY ANTIGEN,
FT FT GOGO-B0101 ALPHA CHAIN.
FT FT EXTRACELLULAR ALPHA-1.
FT FT EXTRACELLULAR ALPHA-2.
FT FT EXTRACELLULAR ALPHA-3.
FT FT CONNECTING PEPTIDE.
FT FT CYTOPLASMIC TAIL.
FT FT BY SIMILARITY.
FT FT BY SIMILARITY.
FT FT BY SIMILARITY.
FT FT BY SIMILARITY.
SQ SEQUENCE 362 AA; 40170 MW; 2E33E2B8 CRC32;

Query Match 44.4%; Score 44; DB 1; Length 362;
Best Local Similarity 90.0%; Pred. No. 5.5;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 11 RENVLRALRY 20
Db 99 RENVLRALRY 108
||||| |||||

RESULT 10
ID B02_GORGO STANDARD; PRT; 362 AA.
AC P30380;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-APR-1993 (Rel. 25, Last annotation update)
DE CLASS I HISTOCOMPATIBILITY ANTIGEN, GOGO-B0102 ALPHA CHAIN PRECURSOR.
OS Gorilla gorilla gorilla (Lowland gorilla)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
RN [1]
RN SEQUENCE FROM N.A.
RP MEDLINE: 92078860.
RX LAWLOR D.A., WARREN E., TAYLOR P., PARHAM P.:

```

CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL; X60254; CAA42806.1; -  
 DR PIR; JH0541; JH0541.  
 DR HSP; P03989; ILSA.  
 DR PROSITE; PS00290; IG\_MHC; 1.  
 DR PFAM; PF00047; IG; 1.  
 DR PFAM; PF00129; MHC\_I; 1.  
 KW MHC I; Transmembrane; Glycoprotein; Signal.  
 FT SIGNAL 1 24  
 FT CHAIN 25 362  
 FT DOMAIN 25 114  
 FT DOMAIN 115 206  
 FT DOMAIN 207 298  
 FT DOMAIN 299 308  
 FT TRANSMEM 309 332  
 FT DOMAIN 333 362  
 FT DISULFID 125 188  
 FT CARBOHYD 227 283  
 FT SEQUENCE 362 AA; 40248 MW; FEAGA941 CRC32;  
 SQ

Query Match 44.4%; Score 44; DB 1; Length 362;  
 Best Local Similarity 90.0%; Pred. No. 5.5;  
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 11 RENLRTALRY 20  
 Db 99 RENLRTALRY 108

RESULT 12  
 1B15\_HUMAN  
 ID 1B15\_HUMAN STANDARD; PRT; 362 AA.  
 AC P10317;  
 DT 01-MAR-1989 (Rel. 10, Created)  
 DT 01-MAR-1989 (Rel. 10, Last sequence update)  
 DT 01-APR-1993 (Rel. 25, Last annotation update)  
 DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-27 B\*2702 ALPHA CHAIN  
 DE PRECURSOR (B-27K) (B27.2).  
 GN HLA-B OR HLAB.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 [1]  
 RP SEQUENCE FROM N.A.  
 RA SEEMANN G.H.A., REIN R.S., BROWN C.S., PLOEGH H.L.;  
 RT "Gene conversion-like mechanisms may generate polymorphism in human  
 RT class I genes.";  
 RL EMBO J. 5:547-552(1986).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA PARHAM P., ARNETT K.L., ADAMS E.J.;  
 RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.  
 [3]  
 RP SEQUENCE OF 86-107 AND 171-181.  
 RX MEDLINE; 86042671.  
 RA LOPEZ DE CASTRO J.A.;  
 RA VEGA M.A., EZQUERRA A., ROJO S., APARICIO P., BRAGADO R.,  
 RT "Structural analysis of an HLA-B27 functional variant: identification  
 RT of residues that contribute to the specificity of recognition by  
 RT cytolytic T lymphocytes.";  
 RL Proc Natl. Acad. Sci. U.S.A. 82:7394-7398(1985).  
 CC -1- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO  
 CC THE IMMUNE SYSTEM.  
 CC -1- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
 CC MICROGLOBULIN).

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL; X03664; CAA27301.1; -  
 DR EMBL; X03667; CAA27301.1; JOINED.  
 DR EMBL; L38504; AAA69724.1; -  
 DR PIR; B25092; HLHUBK.  
 DR HSP; P03989; ILSA.  
 DR MIM; 142830;  
 DR PROSITE; PS00290; IG\_MHC; 1.  
 DR PFAM; PF00047; IG; 1.  
 DR PFAM; PF00129; MHC\_I; 1.  
 KW MHC I; Transmembrane; Glycoprotein; Signal.  
 FT SIGNAL 1 24  
 FT CHAIN 25 362  
 FT DOMAIN 25 114  
 FT DOMAIN 115 206  
 FT DOMAIN 207 298  
 FT DOMAIN 299 308  
 FT TRANSMEM 309 332  
 FT DOMAIN 333 362  
 FT CARBOHYD 110 110  
 FT DISULFID 125 188  
 FT DISULFID 227 283  
 FT SEQUENCE 362 AA; 40397 MW; 9798F0BB CRC32;  
 SQ

Query Match 44.4%; Score 44; DB 1; Length 362;  
 Best Local Similarity 90.0%; Pred. No. 5.5;  
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 11 RENLRTALRY 20  
 Db 99 RENLRTALRY 108

RESULT 13  
 1B47\_HUMAN  
 ID 1B47\_HUMAN STANDARD; PRT; 362 AA.  
 AC P30487;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 01-FEB-1996 (Rel. 33, Last annotation update)  
 DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-49(B-21) B\*4901 ALPHA CHAIN  
 DE PRECURSOR.  
 GN HLA-B OR HLAB.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 [1]  
 RP SEQUENCE FROM N.A.  
 RA PARHAM P., LAWJOR D.A., LOMEN C.E., ENNIS P.D.;  
 RT "Diversity and diversification of HLA-A,B,C alleles.";  
 RL J. Immunol. 142:3937-3950(1989).  
 RN [2]  
 RP REVISION TO 78.  
 RX MEDLINE; 93056529.  
 RA HILDEBRAND W.H., MADRIGAL J.A., BELICH M.P., ZEMMOUR J., WARD F.E.,  
 RA WILLIAMS R.C., PARHAM P.;  
 RT "Serologic cross-reactivities poorly reflect allelic relationships in  
 RT the HLA-B12 and HLA-B21 groups. Dominant epitopes of the alpha 2  
 RT helix.";  
 RL J. Immunol. 149:3563-3568(1992).  
 CC -1- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO  
 CC THE IMMUNE SYSTEM.



CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
 CC MICROGLOBULIN).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; M24037; AAA02950.1; -  
 CC HSSP; P30491; IALM.  
 CC MIM; 142830; -  
 CC PROSITE; PS00290; IG\_MHC; 1.  
 CC PFAM; PF00047; ig; 1.  
 CC PFAM; PF00129; MHC\_I; 1.  
 CC MHC I; Transmembrane; Glycoprotein; Signal.  
 CC SIGNAL 1 24  
 CC CHAIN 25 362  
 CC HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,  
 CC B-49(B-21) B\*4901 ALPHA CHAIN.  
 CC EXTRACELLULAR ALPHA-1.  
 CC DOMAIN 25 114  
 CC DOMAIN 115 206  
 CC DOMAIN 207 298  
 CC DOMAIN 299 309  
 CC TRANSMEM 310 333  
 CC DOMAIN 334 362  
 CC CARBOHYD 110 110  
 CC BY SIMILARITY.  
 CC DISULFID 125 188  
 CC BY SIMILARITY.  
 CC DISULFID 227 283  
 CC BY SIMILARITY.  
 CC SEQUENCE 362 AA; 40581 MW; E996F82F CRC32;

Query Match 44.4%; Score 44; DB 1; Length 362;  
 Best Local Similarity 90.0%; Pred. No. 5.5;  
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 11 RENLTAALRY 20  
 Db 99 RENLRIALRY 108

RESULT 14  
 ID 1B49\_HUMAN STANDARD; PRT; 362 AA.  
 AC P18464;  
 DT 01-NOV-1990 (Rel. 16, Created)  
 DT 01-NOV-1990 (Rel. 16, Last sequence update)  
 DT 01-FEB-1996 (Rel. 33, Last annotation update)  
 DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-51(B-5) B\*5101 ALPHA CHAIN  
 DE PRECURSOR.  
 GN HLA-B OR HLAB.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 90207291.  
 RA ENNIS P.D., ZEMMOUR J., SALTER R.D., PARHAM P.;  
 RT "Rapid cloning of HLA-A,B CDNA by using the polymerase chain  
 RT reaction: frequency and nature of errors produced in amplification.";   
 RL proc. Natl. Acad. Sci. U.S.A. 87:2833-2837(1990).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 89080265.  
 RA HAYASHI H., ENNIS P.D., ARIGA H., SALTER R.D., PARHAM P., KANO K.,  
 RA TAKIGUCHI M.;  
 RT "HLA-B\*51 and HLA-B\*52 differ by only two amino acids which are in the  
 RT helical region of the alpha 1 domain.";   
 RL J. Immunol. 142:306-311(1989).  
 [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 89233295.

RA POHLA H., KUON W., TABACZEWSKI P., DOERNER C., WEISS E.H.;  
 RT "Allelic variation in HLA-B and HLA-C sequences and the evolution of  
 RT the HLA-B alleles".  
 RL Immunogenetics 29:297-307(1989).  
 CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO  
 CC THE IMMUNE SYSTEM.  
 CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
 CC MICROGLOBULIN).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; M2319; AAA36232.1; -  
 DR EMBL; M2792; AAA59620.1; ALT\_SEQ.  
 DR EMBL; M2786; AAA59620.1; JOINED.  
 DR EMBL; M2787; AAA59620.1; JOINED.  
 DR EMBL; M2788; AAA59620.1; JOINED.  
 DR EMBL; M2789; AAA59620.1; JOINED.  
 DR EMBL; M2790; AAA59620.1; JOINED.  
 DR EMBL; M2791; AAA59620.1; JOINED.  
 DR EMBL; L41087; AAA64513.1; -  
 DR EMBL; L41086; AAA64513.1; JOINED.  
 DR PIR; A30345; A30345.  
 DR PIR; A30548; A30548.  
 DR HSSP; P30491; IALM.  
 DR MIM; 142830; -  
 DR PROSITE; PS00290; IG\_MHC; 1.  
 DR PFAM; PF00047; ig; 1.  
 DR PFAM; PF00129; MHC\_I; 1.  
 KW MHC I; Transmembrane; Glycoprotein; Signal.  
 FT SIGNAL 1 24  
 FT CHAIN 25 362

FT HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,  
 FT B-51(B-5) B\*5101 ALPHA CHAIN.  
 FT EXTRACELLULAR ALPHA-1.  
 FT EXTRACELLULAR ALPHA-2.  
 FT EXTRACELLULAR ALPHA-3.  
 FT CONNECTING PEPTIDE.  
 FT CYTOPLASMIC TAIL.  
 FT BY SIMILARITY.  
 FT BY SIMILARITY.  
 FT BY SIMILARITY.  
 FT DISULFID 125 188  
 FT DISULFID 227 283  
 FT BY SIMILARITY.  
 SQ SEQUENCE 362 AA; 40566 MW; 4D846F30 CRC32;

Query Match 44.4%; Score 44; DB 1; Length 362;  
 Best Local Similarity 90.0%; Pred. No. 5.5;  
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 11 RENLTAALRY 20  
 Db 99 RENLRIALRY 108

RESULT 15  
 ID 1B52\_HUMAN STANDARD; PRT; 362 AA.  
 AC P30489;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 01-APR-1993 (Rel. 25, Last annotation update)  
 DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-51(B-5) B\*5104 ALPHA CHAIN  
 DE PRECURSOR.  
 GN HLA-B OR HLAB.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 [1]  
 RN

```

RP SEQUENCE FROM N.A.
RX MEDLINE; 92269955.
RA BELICH M.P., MADRICAL J.A., HILDEBRAND W.H., ZEMMOUR J.,
RA WILLIAMS R.C., LUZ R., PETZL-ERLER M.L., FARHAM P.;
RT "Unusual HLA-B alleles in two tribes of Brazilian Indians.";
RL Nature 357:326-329(1992).
CC -1- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC THE IMMUNE SYSTEM.
CC -1- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC MICROGLOBULIN).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z15143; CAA78849.1; -
CC HSP; P30491; IALM.
CC MM; 142830; ...
CC PROSITE; PS00290; IG_MHC; 1.
CC PFAM; PF00047; Ig; 1.
CC PFAM; PF00129; MHC_I; 1.
CC MHC I; Transmembrane; Glycoprotein; Signal.
CC SIGNAL 1 24
CC CHAIN 25 362 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
CC B-51(B-5) B*5104 ALPHA CHAIN.
CC DOMAIN 25 114 EXTRACELLULAR ALPHA-1.
CC DOMAIN 115 206 EXTRACELLULAR ALPHA-2.
CC DOMAIN 207 298 EXTRACELLULAR ALPHA-3.
CC DOMAIN 299 308 CONNECTING PEPTIDE.
CC TRANSMEM 309 332 CYTOPLASMIC TAIL.
CC DOMAIN 333 362 BY SIMILARITY.
CC CARBOHYD 110 110 BY SIMILARITY.
CC DISULFID 125 188 BY SIMILARITY.
CC DISULFID 227 283 BY SIMILARITY.
CC SEQUENCE 362 AA; 40560 MW; F22F08AB CRC32;
SQ

```

```

Query Match 44.4%; Score 44; DB 1; Length 362;
Best Local Similarity 90.0%; Pred. No. 5.5;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 11 RENDRLALRY 20
Db 99 RENDRLALRY 108

```

```

Search completed: February 8, 2000, 01:26:00
Job time: 1560 sec

```

Result No.	Score	Query			DB	ID	Description
		Match	Length				
1	50	50.5	89	7	O19674	O19674 homo sapien	
2	50	50.5	89	7	O19565	O19565 homo sapien	
3	50	50.5	131	7	O97998	O97998 homo sapien	
4	50	50.5	131	7	O97999	O97999 homo sapien	
5	50	50.5	181	7	O19779	O19779 homo sapien	
6	50	50.5	181	7	O30197	O30197 homo sapien	
7	50	50.5	181	7	O19669	O19669 homo sapien	
8	50	50.5	181	7	O78028	O78028 homo sapien	
9	50	50.5	355	7	O98853	O98853 homo sapien	
10	50	50.5	361	7	O9XR10	O9XR10 pongo pygma	
11	50	50.5	362	7	O29637	O29637 homo sapien	
12	50	50.5	362	7	O29935	O29935 homo sapien	
13	50	50.5	362	7	P79524	P79524 homo sapien	
14	50	50.5	362	7	O29850	O29850 homo sapien	
15	50	50.5	362	7	O29661	O29661 homo sapien	
16	50	50.5	362	7	O78180	O78180 homo sapien	
17	50	50.5	362	7	O29933	O29933 homo sapien	
18	46	46.5	302	5	O25608	O25608 onchocerca	
...19	46	46.5	346	7	O95459	O95459 rattus norv	
20	46	46.5	346	7	O78088	O78088 rattus norv	

Oy	11	RENLTALRY 20
Ddb	74	RENLTALRY 83
	RESULT 2	
	OI9565	
ID	OI9565	PRELIMINARY; PRT; 89 AA.
TAC	OI9565;	
DT	01-JAN-1998	(TREMELrel. 05, Created)
DDT	01-MAY-1999	(TREMELrel. 10, Last sequence update)
DDT	01-MAY-1999	(TREMELrel. 10, Last annotation update)
DE	MHC CLASS I ANTIGEN	(FRAGMENT).
HLA-B.	Homo sapiens	(Human).
OS	Homo sapiens	(Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA CAO K., BURDETT L., ZHANG G., FERNANDEZ-VINA M.;  
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF017316; AAB70282.2; -.  
 KW MHC.  
 FT NON\_TER 1 1  
 FT NON\_TER 89 89  
 SQ SEQUENCE 89 AA; 10581 MW; 9AC7631C CRC32;

Query Match 50.5%; Score 50; DB 7; Length 89;  
 Best Local Similarity 100.0%; Pred. No. 0.8; Indels 0; Gaps 0;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RENLRTALRY 20  
 Db 74 RENLRTALRY 83

RESULT 3  
 O97998  
 ID O97998 PRELIMINARY; PRT; 131 AA.  
 AC O97998;  
 DT 01-MAY-1999 (TRENBLrel. 10, Created)  
 DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TRENBLrel. 10, Last annotation update)  
 DE MHC CLASS I ANTIGEN (FRAGMENT).  
 GN HLA-B.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA HOLDSWORTH R.;  
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF035648; AAD02035.1; -.  
 KW MHC.  
 FT NON\_TER 1 1  
 FT NON\_TER 131 131  
 SQ SEQUENCE 131 AA; 15134 MW; 931C8D27 CRC32;

Query Match 50.5%; Score 50; DB 7; Length 131;  
 Best Local Similarity 100.0%; Pred. No. 1.2; Indels 0; Gaps 0;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RENLRTALRY 20  
 Db 40 RENLRTALRY 49

RESULT 4  
 O97999  
 ID O97999 PRELIMINARY; PRT; 131 AA.  
 AC O97999;  
 DT 01-MAY-1999 (TRENBLrel. 10, Created)  
 DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TRENBLrel. 10, Last annotation update)  
 DE MHC CLASS I ANTIGEN (FRAGMENT).  
 GN HLA-B.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA HOLDSWORTH R.;  
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF035649; AAD02036.1; -.  
 KW MHC.  
 FT NON\_TER 1 1

FT NON\_TER 131 131  
 SQ SEQUENCE 131 AA; 15057 MW; 32C865A3 CRC32;

Query Match 50.5%; Score 50; DB 7; Length 131;  
 Best Local Similarity 100.0%; Pred. No. 1.2;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RENLRTALRY 20  
 Db 40 RENLRTALRY 49

RESULT 5  
 O19779  
 ID O19779 PRELIMINARY; PRT; 181 AA.  
 AC O19779;  
 DT 01-JAN-1998 (TRENBLrel. 05, Created)  
 DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)  
 DT 01-NOV-1999 (TRENBLrel. 12, Last annotation update)  
 DE MHC CLASS I ANTIGEN (FRAGMENT).  
 GN HLA-B.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STEINER N.K., NG J., BUSH J., HARTZMAN R.J., JOHNSON-DOW L.,  
 RA HURLEY C.K.;  
 RL Hum. Immunol. 56:0-0(1997).  
 DR EMBL; U90241; AAB82305.1; -.  
 DR EMBL; U90240; AAB82305.1; JOINED.  
 DR HSP; P10318; IROG.  
 DR PFAM; PF00129; MHC\_I; 1.  
 KW MHC.  
 FT NON\_TER 1 1  
 FT NON\_TER 181 181  
 SQ SEQUENCE 181 AA; 21095 MW; 97EC2597 CRC32;

Query Match 50.5%; Score 50; DB 7; Length 181;  
 Best Local Similarity 100.0%; Pred. No. 1.7;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RENLRTALRY 20  
 Db 74 RENLRTALRY 83

RESULT 6  
 Q30197  
 ID Q30197 PRELIMINARY; PRT; 181 AA.  
 AC Q30197;  
 DT 01-NOV-1996 (TRENBLrel. 01, Created)  
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
 DT 01-NOV-1999 (TRENBLrel. 12, Last annotation update)  
 DE MHC CLASS I ANTIGEN (FRAGMENT).  
 GN HLA-B.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA HURLEY C.K., HOYER R.J.;  
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U63560; AAB05925.1; -.  
 DR EMBL; U63559; AAB05925.1; JOINED.  
 DR HSP; P10318; IROG.  
 DR PFAM; PF00129; MHC\_I; 1.  
 KW MHC.  
 FT NON\_TER 1 1  
 FT NON\_TER 181 181  
 SQ SEQUENCE 181 AA; 20997 MW; 7DFABE37 CRC32;

```

Query Match          50.5%; Score 50; DB 7; Length 181;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RENLRTALRY 20
DB 74 RENLRTALRY 83
|||||

RESULT 7
Q19669 ID O19669 PRELIMINARY; PRT; 181 AA.
AC O19669:
DT 01-JAN-1998 (TREMELREL. 05, Created)
DT 01-JAN-1998 (TREMELREL. 05, Last sequence update)
DT 01-NOV-1999 (TREMELREL. 12, Last annotation update)
DE MHC CLASS I ANTIGEN SHCHA (FRAGMENT).
GN HLA-B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RA HURLEY C.K., STEINER N.K.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U58470; AAB02607.1; -.
DR EMBL; U58469; AAB02607.1; JOINED.
DR HSSP; P10318; 1ROG.
DR PFAM; PF00129; MHC_I; 1.
KW MHC.
FT NON_TER 1
FT NON_TER 181
SQ SEQUENCE 181 AA; 21028 MW; EC872642 CRC32;

Query Match          50.5%; Score 50; DB 7; Length 181;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RENLRTALRY 20
DB 74 RENLRTALRY 83
|||||

RESULT 8
O78028 ID O78028 PRELIMINARY; PRT; 181 AA.
AC O78028:
DT 01-NOV-1998 (TREMELREL. 08, Created)
DT 01-NOV-1998 (TREMELREL. 08, Last sequence update)
DT 01-NOV-1999 (TREMELREL. 12, Last annotation update)
DE HLA-B*15MD.
GN HLA-B*15MD.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RA GAO X., MATHESON B.;
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U58316; AAB87723.1; -.
DR EMBL; U58315; AAB87723.1; JOINED.
DR HSSP; P10318; 1ROG.
DR PFAM; PF00129; MHC_I; 1.
KW MHC.
FT NON_TER 1
FT NON_TER 181
SQ SEQUENCE 181 AA; 21029 MW; C35A18BE CRC32;

Query Match          50.5%; Score 50; DB 7; Length 181;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RENLRTALRY 20
DB 74 RENLRTALRY 83
|||||

RESULT 9
Q29853 ID Q29853 PRELIMINARY; PRT; 355 AA.
AC Q29853:
DT 01-NOV-1996 (TREMELREL. 01, Created)
DT 01-NOV-1996 (TREMELREL. 01, Last sequence update)
DT 01-NOV-1999 (TREMELREL. 12, Last annotation update)
DE HLA-B ALPHA-CHAIN (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RA GAUCHAT-FEISS D., BREUR-VRISENDORP B.S., RUFER N., JEANNET M.,
RA ROOSNEK E., TIERCY J.M.;
RL "Sequencing of a novel functional HLA-B*44 subtype differing in two
RL residues in the alpha 2 domain.";
RL Tissue Antigens 44:261-264(1994).
DR EMBL; X75953; CAA53566.1; -.
DR HSSP; P30491; 1A1M.
DR PROSITE; PS00290; IG_MHC; 1.
DR PFAM; PF00047; Ig; 1.
DR PFAM; PF00129; MHC_I; 1.
KW MHC.
FT NON_TER 355
FT NON_TER 355
SQ SEQUENCE 355 AA; 39845 MW; EF81934E CRC32;

Query Match          50.5%; Score 50; DB 7; Length 355;
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RENLRTALRY 20
DB 99 RENLRTALRY 108
|||||

RESULT 10
Q9XRY0 ID Q9XRY0 PRELIMINARY; PRT; 361 AA.
AC Q9XRY0:
DT 01-NOV-1999 (TREMELREL. 12, Created)
DT 01-NOV-1999 (TREMELREL. 12, Last sequence update)
DT 01-NOV-1999 (TREMELREL. 12, Last annotation update)
DE MHC CLASS I ANTIGEN (FRAGMENT).
GN POPY B.
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Pongo.
RN [1]
RP SEQUENCE FROM N.A.
RA ADAMS E.J., THOMSON G., PARHAM P.;
RL "Evidence for an HLA-C-like locus in the orangutan Pongo pygmaeus.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF118895; AAD28435.1; -.
DR PROSITE; PS00290; IG_MHC; 1.
FT NON_TER 1
FT NON_TER 361
SQ SEQUENCE 361 AA; 40466 MW; CB177B4F CRC32;

Query Match          50.5%; Score 50; DB 7; Length 361;
Best Local Similarity 100.0%; Pred. No. 3.4;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 11 RENLRTALRY 20  
Db 98 RENLRTALRY 107  
RESULT 11  
Q29637 PRELIMINARY; PRT; 362 AA.  
ID Q29637;  
AC Q29637;  
DT 01-NOV-1996 (TRENBLrel. 01, Created)  
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
DT 01-NOV-1999 (TRENBLrel. 12, Last annotation update)  
DE MHC CLASS I ANTIGEN.  
GN HLA-B.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA DOMENA J.D., PARHAM P.;  
RL Submitted (DEC-1993) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U04244; AA87397.1; -.  
DR HSSP; P30460; IALM.  
DR PROSITE; PS00290; IG\_MHC; 1.  
DR PFAM; PF00047; Ig; 1.  
DR PFAM; PF00129; MHC\_I; 1.  
KW MHC.  
SQ SEQUENCE 362 AA; 40359 MW; E8AB95D7 CRC32;

Query Match 50.5%; Score 50; DB 7; Length 362;  
Best Local Similarity 100.0%; Pred. No. 3.4;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RENLRTALRY 20  
Db 99 RENLRTALRY 108  
RESULT 12  
Q29935 PRELIMINARY; PRT; 362 AA.  
ID Q29935;  
AC Q29935;  
DT 01-NOV-1996 (TRENBLrel. 01, Created)  
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
DT 01-NOV-1999 (TRENBLrel. 12, Last annotation update)  
DE MHC CLASS I HLA-B\*3802.  
GN HLA-B.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA LITTLE A.M.M., DOMENA J.D., HILDEBRAND W.H., PARHAM P.;  
RL Submitted (FEB-1994) to the EMBL/GenBank/DBJ databases.  
DR EMBL; L22028; AAA59618.1; -.  
DR HSSP; P30491; IALM.  
DR PROSITE; PS00290; IG\_MHC; 1.  
DR PFAM; PF00047; Ig; 1.  
DR PFAM; PF00129; MHC\_I; 1.  
KW MHC.  
SQ SEQUENCE 362 AA; 40404 MW; 83A8399C CRC32;

Query Match 50.5%; Score 50; DB 7; Length 362;  
Best Local Similarity 100.0%; Pred. No. 3.4;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RENLRTALRY 20  
Db 99 RENLRTALRY 108

RESULT 13  
P79524 PRELIMINARY; PRT; 362 AA.  
ID P79524;  
AC P79524;  
DT 01-MAY-1997 (TRENBLrel. 03, Created)  
DT 01-MAY-1997 (TRENBLrel. 03, Last sequence update)  
DT 01-NOV-1999 (TRENBLrel. 12, Last annotation update)  
DE MHC CLASS I HLA-B\*440.  
GN HLA-B.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA ARNETT K.L., DARKE C., PARHAM P.;  
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U64801; AAB40632.1; -.  
DR HSSP; P30491; IALM.  
DR PROSITE; PS00290; IG\_MHC; 1.  
DR PFAM; PF00047; Ig; 1.  
DR PFAM; PF00129; MHC\_I; 1.  
KW MHC.  
SQ SEQUENCE 362 AA; 40396 MW; AC0FAF3C CRC32;

\* Query Match 50.5%; Score 50; DB 7; Length 362;  
Best Local Similarity 100.0%; Pred. No. 3.4;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RENLRTALRY 20  
Db 99 RENLRTALRY 108

RESULT 14  
Q29850 PRELIMINARY; PRT; 362 AA.  
ID Q29850;  
AC Q29850;  
DT 01-NOV-1996 (TRENBLrel. 01, Created)  
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
DT 01-NOV-1999 (TRENBLrel. 12, Last annotation update)  
DE HLA-B ALPHA CHAIN B\*4407 PRECURSOR.  
GN HLA-B\*4407.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 97004417.  
RA VILCHES C., SANZ L., DE PABLO R., MORENO M.E., PUENTE S., KREISLER M.;  
RT "Molecular characterization of the new alleles HLA-B\*8101 and B\*4407.";  
RL Tissue Antigens 47:139-142(1996).  
DR EMBL; X90391; CAA62036.1; -.  
DR HSSP; P30491; IALM.  
DR PROSITE; PS00290; IG\_MHC; 1.  
DR PFAM; PF00047; Ig; 1.  
DR PFAM; PF00129; MHC\_I; 1.  
KW Signal; MHC.  
FT SIGNAL 1 24 POTENTIAL.  
FT CHAIN 25 362 HLA-B ALPHA CHAIN B\*4407.  
SQ SEQUENCE 362 AA; 40449 MW; 07B5245D CRC32;

Query Match 50.5%; Score 50; DB 7; Length 362;  
Best Local Similarity 100.0%; Pred. No. 3.4;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RENLRTALRY 20  
Db 99 RENLRTALRY 108

```

RESULT 15
Q29661
ID Q29661 PRELIMINARY; PRT; 362 AA.
AC Q29661;
DT 01-NOV-1996 (TREMELREL. 01, Created)
DT 01-NOV-1996 (TREMELREL. 01, Last sequence update)
DT 01-NOV-1999 (TREMELREL. 12, Last annotation update)
DE MHC ANTIGEN HLA-B*1303 PRECURSOR.
GN HLA-B*1303.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BLOOD;
RX MEDLINE: 96232969.
RA BALAS A., GARCIA-SANCHEZ F., VICARIO J.L.;
RT "HLA-B*1303: a new example of poor correlation between serology and
RT structure.";
RL Hum. Immunol. 45:32-36(1996).
DR EMBL: U14943; AAB06829.1; -.
DR HSP; P30491; IAIW.
DR PROSITE; PS00290; IG_MHC; 1.
DR PFAM; PF00047; Ig; 1.
DR PFAM; PF00129; MHC_I; 1.
KW Signal; MHC.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 362 MHC ANTIGEN HLA-B*1303.
SQ SEQUENCE 362 AA; 40501 MW; 9277DEFB CRC32;

```

```

Query Match 50.5%; Score 50; DB 7; Length 362;
Best Local Similarity 100.0%; Pred. No. 3.4;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 11 RENLRTALRY 20
Db 99 RENLRTALRY 108
|||||

```

Search completed: February 8, 2000, 19:16:14  
Job time: 21501 sec

**THIS PAGE BLANK (USPTO)**



Sequence	Strd Orig	2Score	EScore	Len	Documentation	...	
gb_hg9:AC005423	+	51.00	97.81	127934	!	AC068423 Arabidopsis thalian	
gb_hg9:AC011984	+	51.00	95.76	160437	!	AC081984 Homo sapiens clone	
gb_pr1:HSB44HLA2	+	50.00	154.99	1.31	163	!	X78849 H.sapiens HLA-B gene, ex
gb_pr1:HSHLAB2	+	50.00	153.75	1.53	187	!	X78426 H.sapiens HLA-B gene, ex
gb_pr1:HOSA18648	+	50.00	150.42	2.35	270	!	X18648 Homo sapiens HLA-B gene,
gb_pr1:HSAL32659	+	50.00	150.42	2.35	270	!	X132659 Homo sapiens HLA-B gene,
gb_pr1:HSHLAB132	+	50.00	150.42	2.35	270	!	X12378 H.sapiens HLA-B*13 gene,
gb_pr2:HSAC49724	+	50.00	150.42	2.35	270	!	AJ49724 Homo sapiens partial H
gb_pr2:HSBHC28701	+	50.00	150.42	2.35	270	!	X58469 Human MHC class I anti
gb_pr2:HSBHC49121	+	50.00	150.42	2.35	270	!	U58413 Human HLA-B*15MD gene, e
gb_pr3:HSHLABSH1	+	50.00	150.42	2.35	270	!	U59313 Human MHC class I anti
gb_pr3:HSHLABSH1	+	50.00	150.42	2.35	270	!	U90240 Human MHC class I anti
gb_pr4:AF017316	+	50.00	150.42	2.35	270	!	AF017316 Homo sapiens MHC class
gb_pr4:HB38021G1	+	50.00	150.42	2.35	270	!	AF181857 Homo sapiens HLA-B MHC
gb_pr4:HS44032S1	+	50.00	150.42	2.35	270	!	U63559 Human MHC class I anti
gb_pr4:AF035648	+	50.00	146.97	3.66	395	!	AF035648 Homo sapiens MHC class
gb_pr4:AF035649	+	50.00	146.97	3.66	395	!	AF035649 Homo sapiens MHC class
gb_pr2:HUMBW44HLA	+	50.00	146.77	3.75	404	!	L31798 Homo sapiens HLA-B*W44 va
gb_pr2:AB032598	+	50.00	144.30	5.15	531	!	AB032598 Homo sapiens HLA-B1301
gb_pr2:HSAL33267	+	50.00	144.04	5.32	546	!	AJ133267 Homo sapiens mRNA for
gb_pr2:HMB2270	+	50.00	144.04	5.32	546	!	L76933 Homo sapiens MHC class I
gb_pr2:HSAL238702	+	50.00	139.34	9.73	918	!	AJ238702 Homo sapiens MHC class
gb_pr4:AF056981	+	50.00	139.06	10.09	947	!	AF056981 Homo sapiens MHC class
gb_pr4:U90561	+	50.00	139.06	10.09	947	!	U90561 Homo sapiens MHC class I
gb_pr1:HSHLABAB	+	50.00	138.00	11.56	1064	!	X75953 H.sapiens HLA-B mRNA, cl
gb_pr4:AF118995	+	50.00	137.79	11.87	1088	!	AF118995 Pongo pygmaeus MHC cl
gb_pr1:HSHLAB44	+	50.00	137.79	11.87	1089	!	X64366 H.sapiens HLA-B*4403 mR
gb_pr1:HSHLAB484G	+	50.00	137.79	11.87	1089	!	X90391 H.sapiens mRNA for HLA-B
gb_pr1:HUMB1301A	+	50.00	137.79	11.87	1089	!	D50290 Human mRNA for HLA-B*13
gb_pr1:HUMB1302B	+	50.00	137.79	11.87	1089	!	D50291 Human mRNA for HLA-B*13
gb_pr1:HUMBHB38A	+	50.00	137.79	11.87	1089	!	L22028 Human MHC class I HLA-E
gb_pr1:HUMMHCB13A	+	50.00	137.79	11.87	1089	!	M24041 Human MHC class I HLA-E
gb_pr1:HUMMHCB44A	+	50.00	137.79	11.87	1089	!	M24038 Human MHC class I HLA-E
gb_pr2:HSAL31118	+	50.00	137.79	11.87	1089	!	AJ131118 Homo sapiens mRNA for
gb_pr2:HSU04244	+	50.00	137.79	11.87	1089	!	U04244 Human MHC class I anti
gb_pr2:HSU064801	+	50.00	137.79	11.87	1089	!	U64801 Human MHC class I HLA-B
gb_pr2:HUMHLABK	+	50.00	137.79	11.87	1089	!	L42282 Homo sapiens (clone 18)
gb_pr2:HUMHLABL	+	50.00	137.79	11.87	1089	!	L42283 Homo sapiens major hist
gb_pr3:HSU75533	+	50.00	137.79	11.87	1089	!	U75533 Homo sapiens HLA class
gb_pr1:HUMHLAB13	+	50.00	137.76	11.92	1093	!	M19757 Human MHC class I HLA-E
gb_pr1:HUMHMB13A	+	50.00	137.76	11.92	1093	!	M24075 Human MHC class I HLA-E
gb_pr2:HSU14943	+	50.00	136.18	14.60	1301	!	U14943 Human MHC antigen (HLA-
gb_in1:CEVITELA	+	50.00	123.02	78.89	5562	!	X56212 C.elegans gene for vitel

US-08-653-294-29 x AC006423 ..

US-08-653-294-29 x AC006423 ::



```

FEATURES
  source
    Location/Qualifiers
      1..163
        /organism="Homo sapiens"
        /isolate="WJG"
        /db_xref="taxon:9606"
        /chromosome="6"
        /germline
        /note="MHC"
      exon
        1..163
        /note="MHC Class I heavy chain HLA-B44 variant"
        /number=2
  BASE COUNT
    45 a 50 c 54 g 14 t
  ORIGIN

alignment_scores:
  Quality: 50.00 Length: 10
  Ratio: 5.000 Gaps: 0
  Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
  US-08-653-294-29 x HSB44HLA2 ..
  Align seg 1/1 to: HSB44HLA2 from: 1 to: 163

11 ArgGluAsnLeuArgThrAlaLeuArgTyr 20
|||||
115 CGAGAGAACCTGGCACCCTCCGCTAC 144

seq_name: gb_pri:HSB44HLA2

seq_documentation_block:
  LOCUS HSB44HLA2 187 bp DNA PRI 15-APR-1994
  DEFINITION H.sapiens HLA-B gene, exon 2.
  ACCESSION X78426
  VERSION X78426.1 GI:473237
  KEYWORDS heavy chain; HLA-B*4404 gene; MHC class I associated.
  SOURCE human.
  ORGANISM Homo sapiens
    Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
    Primates; Catarrhini; Hominidae; Homo.
  REFERENCE
    1 (bases 1 to 187)
  AUTHORS Yao.Z. and Albert.E.
  TITLE Identification of two major HLA-B44 subtypes an a novel B44
    sequence: oligotyping
  JOURNAL Unpublished
  AUTHORS Yao.Z.
  TITLE Direct Submission
  JOURNAL Submitted (18-MAR-1994) Z. Yao, Immunogenetics Laboratory,
    Kinderpoliklinik University of Munich, Pettenkaferstr 8a, 80336
    Munich, FRG
  REMARK revised by [3] MAT
  REFERENCE 3 (bases 1 to 187)
  AUTHORS Yao.Z.
  TITLE Direct Submission
  JOURNAL Submitted (15-APR-1994) Z. Yao, Immunogenetics Laboratory,
    Kinderpoliklinik University of Munich, Pettenkaferstr 8a, 80336
    Munich, FRG
  COMMENT On Apr 17, 1994 this sequence version replaced gi:468539.
FEATURES
  source
    Location/Qualifiers
      1..187
        /organism="Homo sapiens"
        /isolate="BEB"
        /db_xref="taxon:9606"
        /chromosome="6"
        /cell_line="BEB"
      exon
        1..187
        /gene="HLA-B44"
      exon
        <1..187
        /gene="HLA-B44"
        /number=2
  BASE COUNT
    49 a 56 c 62 g 20 t

FEATURES
  source
    Location/Qualifiers
      1..270
        /organism="Homo sapiens"
        /isolate="individual 22658"
        /db_xref="taxon:9606"
        /rearranged
        /cell_type="leucocyte"
        /tissue_type="blood"
        /cell_line="22658"
      exon
        <1..>270
        /gene="HLA-B"
        /note="B*0802 allele"
        /number=2
      gene
        1..270
        /gene="HLA-B"
        <1..>270
        /gene="HLA-B"
        /standard_name="MHC class I antigen"
        /note="B*0802 allele"
        /codon_start=3
        /product="human leucocyte antigen B alpha chain"
        /protein_id="CAA77241.1"
        /db_xref="GI:4160523"
        /translation="SHSMRYFDTAMSRGEGPRFISGVYDDTQFVRFDSDAASPRE
          EPRAPWIEGEGPEYDNRNTQIFKNTQDRENLRTALRYNOSEA"
      CDS
        60 a 88 c 87 g 35 t
      ORIGIN

alignment_scores:
  Quality: 50.00 Length: 10
  Ratio: 5.000 Gaps: 0
  Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
  US-08-653-294-29 x HSLAB2 ..
  Align seg 1/1 to: HSLAB2 from: 1 to: 187

11 ArgGluAsnLeuArgThrAlaLeuArgTyr 20
|||||
139 CGAGAGAACCTGGCACCCTCCGCTAC 168

seq_name: gb_pri:HOSA18648

seq_documentation_block:
  LOCUS HOSA18648 270 bp DNA PRI 15-JAN-1999
  DEFINITION Homo sapiens HLA-B gene, allele HLA-B*0802, exon 2.
  ACCESSION Y18648
  VERSION Y18648.1 GI:4160522
  KEYWORDS HLA-B gene; HLA-B*0802 allele; human leucocyte antigen B; major
    histocompatibility complex; MHC class I antigen.
  SOURCE human.
  ORGANISM Homo sapiens
    Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
    Primates; Catarrhini; Hominidae; Homo.
  REFERENCE
    1 (bases 1 to 270)
  AUTHORS Guttridge,M.G
  TITLE Direct Submission
  JOURNAL Submitted (12-JAN-1999) M.G. Guttridge, Welsh Blood Service, Ely
    Valley Road, Talbot Green, Pontyclun, CF72 9WB, UK
  FEATURES
    Location/Qualifiers
      1..270
        /organism="Homo sapiens"
        /isolate="individual 22658"
        /db_xref="taxon:9606"
        /rearranged
        /cell_type="leucocyte"
        /tissue_type="blood"
        /cell_line="22658"
      exon
        <1..>270
        /gene="HLA-B"
        /note="B*0802 allele"
        /number=2
      gene
        1..270
        /gene="HLA-B"
        <1..>270
        /gene="HLA-B"
        /standard_name="MHC class I antigen"
        /note="B*0802 allele"
        /codon_start=3
        /product="human leucocyte antigen B alpha chain"
        /protein_id="CAA77241.1"
        /db_xref="GI:4160523"
        /translation="SHSMRYFDTAMSRGEGPRFISGVYDDTQFVRFDSDAASPRE
          EPRAPWIEGEGPEYDNRNTQIFKNTQDRENLRTALRYNOSEA"
      CDS
        60 a 88 c 87 g 35 t
      ORIGIN

alignment_scores:
  Quality: 50.00 Length: 10
  Ratio: 5.000 Gaps: 0
  Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
  US-08-653-294-29 x HOSA18648 ..
  Align seg 1/1 to: HOSA18648 from: 1 to: 270

```

11 ArgGluAsnLeuArgThrAlaLeuArgTyr 20  
 |||  
 222 CGAGAGAACCTGGCCACCGCTCCGCTAC 251

seq\_name: gb\_pr1: HSA132659

seq\_documentation\_block: 270 bp DNA PRI 02-FEB-1999  
 LOCUS HSA132659 Homo sapiens HLA-B gene, B\*408 allele, exon 2 and joined CDS.  
 DEFINITION AJ132659  
 ACCESSION  
 VERSION AJ132659.1 GI:4218239  
 KEYWORDS HLA-B gene; HLA-B\*4408 allele; human leucocyte antigen B; major histocompatibility complex; MHC class I antigen.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;

Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 270)

AUTHORS Guttridge M.G.

TITLE Direct Submission

JOURNAL Submitted (01-FEB-1999) Guttridge M.G., Tissue Typing, Welsh Blood Service, Ely Valley Road, Talbot Green, Pontyclun, Cardiff, CF72

9WB, UNITED KINGDOM

REFERENCE 2 (bases 1 to 270)

AUTHORS Guttridge M.G.

TITLE Serology and confirmatory sequence of HLA-B\*4408

JOURNAL Unpublished

FEATURES Location/Qualifiers

1..270

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/sex="Male"

/cell\_type="lymphocytes"

/cell\_line="ttid 30458"

1..270

/gene="HLA-B"

/note="B\*4408 allele"

/number=2

Join(<1..270,AJ132660.1:1..>276)

/gene="HLA-B"

Join(<1..270,AJ132660.1:1..>276)

/gene="HLA-B"

/standard\_name="MHC class I antigen"

/note="B\*4408 allele"

/codon\_start=3

/label="hlab.cds"

/product="human leucocyte antigen B"

/protein\_id="CAA10726.1"

/db\_xref="GI:4218240"

/translation="SHSMRYFYTAMSRGPRGPRFITVGVYDDTLFVRFSDAASPRM

APRAPWIEQSGPEYWDRETQISKNTQYRNLRLALRYNQSGSHIIQRMIGCDV

GPDRLLRGYDDAYDGDYIADNEDLSSTAAADTAQAQITQRKWEAKVAEQDRAYLE

GLCVESLRYLENGKETLQRA"

BASE COUNT 57 a 91 c 85 g 37 t

ORIGIN

alignment\_scores:  
 Quality: 50.00 Length: 10  
 Ratio: 5.000 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-08-653-294-29 x HSA132659 ..

Align seg 1/1 to: HSA132659 from: 1 to: 270

11 ArgGluAsnLeuArgThrAlaLeuArgTyr 20

|||||

222 CGAGAGAACCTGGCCACCGCTCCGCTAC 251

seq\_name: gb\_pr1: HSHLAB132

seq\_documentation\_block:

LOCUS HSHLAB132 270 bp DNA PRI 28-APR-1997

DEFINITION H.sapiens HLA-B\*13 gene, variant exon 2.

ACCESSION Y12378

VERSION Y12378.1 GI:1934896

KEYWORDS exon 2; HLA-B\*13.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;

Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 270)

AUTHORS Maertens,R. and De Canck,I.

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 270)

AUTHORS De Canck,I.

TITLE Direct Submission

JOURNAL Submitted (07-APR-1997) I. De Canck, INNOGENETICS N.V.,

Industriepark Zwijsnaarde 7 box 4, B-9052 Gent, BELGIUM

FEATURES Location/Qualifiers

1..270

/organism="Homo sapiens"

/isolate="76002"

/db\_xref="taxon:9606"

/chromosome="6"

/map="p"

<1..>270

/gene="HLA-B\*13"

/number=2

1..270

/gene="HLA-B\*13"

<1..>270

/gene="HLA-B\*13"

/codon\_start=3

/protein\_id="CAA73021.1"

/db\_xref="GI:1934897"

/translation="SHSMRYFYTAMSRGPRGPRFITVGVYDDTLFVRFSDAASPRM

APRAPWIEQSGPEYWDRETQISKNTQYRNLRLALRYNQSEA"

BASE COUNT 59 a 92 c 83 g 36 t

ORIGIN

alignment\_scores:  
 Quality: 50.00 Length: 10  
 Ratio: 5.000 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-08-653-294-29 x HSHLAB132 ..

Align seg 1/1 to: HSHLAB132 from: 1 to: 270

11 ArgGluAsnLeuArgThrAlaLeuArgTyr 20

|||||

222 CGAGAGAACCTGGCCACCGCTCCGCTAC 251

seq\_name: gb\_pr2: HSA249724

seq\_documentation\_block:

LOCUS HSA249724 270 bp DNA PRI 06-OCT-1999

DEFINITION Homo sapiens partial HLA-B gene for MHC class I antigen, HLA-B\*44

variant, exon 2.

ACCESSION AJ249724

VERSION AJ249724.1 GI:5919229

KEYWORDS HLA-B gene; HLA-B\*44 variant; human leucocyte antigen B; major histocompatibility complex; MHC class I antigen.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 270)

AUTHORS Elsner,H.A., Schmitz,G. and Blasczyk,R.

```

TITLE      A new HLA-B*44 variant
JOURNAL    Unpublished
REFERENCE  2 (bases 1 to 270)
AUTHORS    Blasczyk,R.
TITLE      Direct Submission
JOURNAL    Medicine, Hannover Medical School, Carl-Neuberg-Str. 1, Hannover,
30625, GERMANY

FEATURES   Location/Qualifiers
source     1..270
            /organism="Homo sapiens"
            /isolate="MH997336/B*1268"
            /db_xref="taxon:9606"
            /chromosome="6"
            /map="6p21"
            1..270
            /gene="HLA-B"
            /note="HLA-B*44 variant"
            /number=2
            1..270
            /gene="HLA-B"
            /gene="HLA-B"
            /standard_name="MHC class I antigen"
            /note="HLA-B*44 variant"
            /codon_start=3
            /label="HLA-B*44_CDS"
            /product="human leucocyte antigen B"
            /protein_id="CAB56342.2"
            /db_xref="GI:6015538"
            /translation="SHSMRYFYTAMSRGPRGEPREITGVDDTLFVRFDSDATSPRK
            EPRAPWIEQEGPEYWDRETQISKNTQTYRENLRALRYNOSEA"

BASE COUNT 60 a 89 c 85 g 36 t
ORIGIN

alignment_scores:
  Quality: 50.00 Length: 10
  Ratio: 5.000 Gaps: 0
  Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
  US-08-653-294-29 x HSSHCHA01 ..
  Align seg 1/1 to: HSSHCHA01 from: 1 to: 270
  11 ArgGluAsnLeuArgThrAlaLeuArgTyr 20
  |||||
  222 CGAGAGAACCTGCGACCGCGCTCCGCTAC 251
  seq_name: gb_pr2:HSSHCHA01

seq_documentation_block:
  LOCUS HSSHCHA01 270 bp DNA PRI 18-JUN-1996
  DEFINITION Human MHC class I antigen SHCHA (HLA-B*4403 variant) gene, exon 2.
  ACCESSION U58469
  VERSION U58469.1 GI:1378136
  KEYWORDS
  SOURCE 1 of 2
  ORGANISM human.
  Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;
  Eutheria; Primates; Catarrhini; Hominidae; Homo.
  REFERENCE 1 (bases 1 to 270)
  AUTHORS Hurley,C.K. and Steiner,N.K.
  TITLE Novel HLA-B alleles
  JOURNAL Unpublished
  REFERENCE 2 (bases 1 to 270)
  AUTHORS Hurley,C.K. and Steiner,N.K.
  TITLE Direct Submission
  JOURNAL Submitted (16-MAY-1996) C.K. Hurley, Microbiology & Immunology,
  Georgetown University, 3970 Reservoir Rd.NW, Washington, DC 20007,
  USA

```

```

FEATURES   Location/Qualifiers
source     1..270
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /cell_type="PBL"
            /haplotype="B*5101/B*4403 variant"
            /chromosome="6"
            /note="SHCHA-B*4403 variant, Direct sequence"
            1..270
            /gene="HLA-B"
            /note="HLA-B*4403 variant"
            /number=2
            60 a 89 c 85 g 36 t
            BASE COUNT
            ORIGIN

alignment_scores:
  Quality: 50.00 Length: 10
  Ratio: 5.000 Gaps: 0
  Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
  US-08-653-294-29 x HSSHCHA01 ..
  Align seg 1/1 to: HSSHCHA01 from: 1 to: 270
  11 ArgGluAsnLeuArgThrAlaLeuArgTyr 20
  |||||
  222 CGAGAGAACCTGCGACCGCGCTCCGCTAC 251
  seq_name: gb_pr3:HSHLAB15M1

seq_documentation_block:
  LOCUS HSHLAB15M1 270 bp DNA PRI 01-DEC-1997
  DEFINITION Human HLA-B*15MD gene, exon 2.
  ACCESSION U58315
  VERSION U58315.1 GI:2654406
  KEYWORDS
  SOURCE 1 of 2
  ORGANISM human.
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
  Eutheria; Primates; Catarrhini; Hominidae; Homo.
  REFERENCE 1 (sites)
  AUTHORS Gao,X. and Matheson,B.
  TITLE A novel B15 variant found in oceanic populations
  JOURNAL Unpublished
  REFERENCE 2 (bases 1 to 270)
  AUTHORS Gao,X.
  TITLE Direct Submission
  JOURNAL Submitted (15-MAY-1996) Xiaojiang Gao, The Australian National
  University, Human Genetics Group, Acton, Canberra, Act, Australia,
  2601

FEATURES   Location/Qualifiers
source     1..270
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="6"
            /map="6P21"
            /cell_line="MD674"
            /cell_type="lymphocyte"
            /tissue_type="blood"
            1..270
            /gene="HLA-B*15MD"
            /number=2
            58 a 91 c 85 g 36 t
            BASE COUNT
            ORIGIN

alignment_scores:
  Quality: 50.00 Length: 10
  Ratio: 5.000 Gaps: 0
  Percent Similarity: 100.000 Percent Identity: 100.000

```

## alignment\_block:

US-08-653-294-29 x HSHLAB15M1 ..

Align seg 1/1 to: HSHLAB15M1 from: 1 to: 270

11 ArgGluAsnLeuArgThrAlaLeuArgTyr 20  
 |||  
 222 CGAGAGAACCTGCGCACCGCTCGCTAC 251

seq\_name: gb\_pr3:HSHLABSH1

seq\_documentation\_block:

LOCUS HSHLABSH1 270 bp DNA PRI 30-OCT-1997  
 DEFINITION Human MHC class I antigen HLA-B GN00155-B\*38022 gene, exon 2.  
 ACCESSION U90240  
 VERSION U90240.1 GI:1905853  
 KEYWORDS  
 SEGMENT 1 of 2  
 SOURCE human.

ORGANISM

Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 270)  
 Steiner,N.K., Ng,J., Bush,J., Hartzman,R.J., Johnson-Dow,L. and  
 Hurley,C.K.

AUTHORS

HLA-B Alleles Associated with the B15 Serologically Defined

TITLE

Antigens

JOURNAL

Hum. Immunol. 56 (1997) In press

REFERENCE

2 (bases 1 to 270)  
 Steiner,N.K. and Hurley,C.K.

AUTHORS

TITLE

Direct Submission

JOURNAL

Submitted (20-FEB-1997) Microbiology & Immunology, Georgetown  
 University, 3970 Reservoir Rd. NW, Washington, DC 20007, USA

FEATURES

source

1. .270

/organism="Homo sapiens"

/isolate="GN00155"

/db\_xref="taxon:9606"

1. .270

/gene="HLA-B"

/number=2

BASE COUNT 60 a 88 c 85 g 37 t

ORIGIN

seq\_name: gb\_pr4:HSHLABSH1

seq\_documentation\_block:

LOCUS AF017316 270 bp DNA PRI 26-MAR-1999  
 DEFINITION Homo sapiens MHC class I antigen (HLA-B) gene, partial cds.  
 ACCESSION AF017316  
 VERSION AF017316.1 GI:2394335  
 KEYWORDS  
 SEGMENT 1 of 2  
 SOURCE human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 270)

## AUTHORS

Cao,K., Burdett,L., Zhang,G. and Fernandez-Vina,M.

## TITLE

Direct Submission

## JOURNAL

Submitted (07-AUG-1997) Nat. Histocompatibility Lab, Am Red Cross,  
 box 173, 22 S. Greene St., Baltimore, MD 21201, USA

## FEATURES

Location/Qualifiers

1. .270

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="27B"

/chromosome="6"

&lt;1. .&gt;270

/gene="HLA-B"

/note="novel allele"

&lt;1. .&gt;270

/gene="HLA-B"

/codon\_start=3

/product="MHC class I antigen"

/protein\_id="AAB70282.2"

/db\_xref="GI:4521330"

/translation="SHSMRYFYTAMSRPGEGEPFITYGVYDDTQFVRFSDATSPRM  
 APRAPWIEGPEYWDRETOISNTQTRENLRTALRYNQSEA"

BASE COUNT 59 a 92 c 83 g 36 t

ORIGIN

## alignment\_scores:

Quality: 50.00 Length: 10

Ratio: 5.000 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-08-653-294-29 x AF017316 ..

Align seg 1/1 to: AF017316 from: 1 to: 270

11 ArgGluAsnLeuArgThrAlaLeuArgTyr 20

|||

222 CGAGAGAACCTGCGCACCGCTCGCTAC 251

seq\_name: gb\_pr4:HB38021G1

seq\_documentation\_block:

LOCUS HB38021G1 270 bp DNA PRI 23-SEP-1999  
 DEFINITION Homo sapiens HLA-B MHC class I antigen (HLA-B) gene, HLA-B\*38021  
 allele, exon 2.

ACCESSION AF181857

VERSION AF181857.1 GI:5919134

KEYWORDS

SEGMENT 1 of 2

SOURCE human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 270)

Pimtanothai,N. and Hurley,C.K.

TITLE

Novel HLA-B allele

JOURNAL

Unpublished

REFERENCE

2 (bases 1 to 270)

Pimtanothai,N. and Hurley,C.K.

TITLE

Direct Submission

JOURNAL

Submitted (30-AUG-1999) Microbiology & Immunology, Georgetown  
 University, 3970 Reservoir Rd. N.W., Washington, D.C. 20007, USA

FEATURES

Location/Qualifiers

1. .270

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="49N-TA"

&lt;1. .&gt;270

/gene="HLA-B"

/number=2

BASE COUNT 58 a 87 c 88 g 37 t

ORIGIN

```

alignment_scores:
  Quality: 50.00      Length: 10
  Ratio: 5.000       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
  US-08-653-294-29 x HB38021G1 ..
  Align seg 1/1 to: HB38021G1 from: 1 to: 270

11 ArgGluAsnLeuArgThrAlaLeuArgTyr 20
|||||
222 CGAGAGAACCTGCGCACCAGCGCTCCGCTAC 251

seq_name: gb_pr4:HS44032S1

seq_documentation_block:
  LOCUS HS44032S1 270 bp DNA PRI 21-JUN-1999
  DEFINITION Human MHC class I antigen HLA-B (HLA-B*4410 allele) gene, exon 2.
  ACCESSION U63559
  VERSION U63559.1 GI:1488303
  KEYWORDS
  SEGMENT 1 of 2
  SOURCE human.
  ORGANISM Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
  Eutheria; Primates; Catarrhini; Hominidae; Homo.
  REFERENCE
  AUTHORS Hurley,C.K. and Hoyer,R.J.
  TITLE Human MHC Class I Antigen HLA-B*4410
  JOURNAL Unpublished
  REFERENCE 2 (bases 1 to 270)
  AUTHORS Hurley,C.K. and Hoyer,R.J.
  TITLE Direct Submission
  JOURNAL Submitted (11-JUL-1996) Dept. of Microbiology and Immunology,
  Georgetown University, 3970 Reservoir Rd., NW, Washington, DC
  20007, USA
  FEATURES
    source
      1..270
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /chromosome="6"
        /cell_type="white blood cells from whole blood"
    exon
      1..270
        /gene="HLA-B"
        /number=2
  BASE COUNT 60 a 89 c 85 g 36 t
  ORIGIN

```

```

KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 395)
AUTHORS Holdsworth,R.
TITLE Direct Submission
JOURNAL Submitted (24-NOV-1997) Tissue Typing, Red Cross Blood Service,
South Melbourne, Melbourne, Victoria 3205, Australia
FEATURES
  Location/Qualifiers
    source
      1..395
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /chromosome="6"
        <1..>395
        /gene="HLA-B"
        /product="MHC class I antigen"
        <1..>395
        /gene="HLA-B"
        /allele="HLA-B*4402 variant"
        <1..>395
        /gene="HLA-B"
        /codon_start=3
        /product="MHC class I antigen"
        /protein_id="AAD02035.1"
        /db_xref="GI:4104423"
        /translation="FSDATSPKPEPRAPWISQEGPEYWDRETOISKNTQTYRENLR
        TALRYNQSEAGSHIIQRMYGDPDGRLLRGYDQDAYDGKDIYALNEDLSSTAAAD
        TAAQITQKWEARVAEODRAYIEGCV"
  BASE COUNT 89 a 120 c 141 g 45 t
  ORIGIN

```

```

alignment_scores:
  Quality: 50.00      Length: 10
  Ratio: 5.000       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
  US-08-653-294-29 x AF035648 ..
  Align seg 1/1 to: AF035648 from: 1 to: 395

11 ArgGluAsnLeuArgThrAlaLeuArgTyr 20
|||||
120 CGAGAGAACCTGCGCACCAGCGCTCCGCTAC 149

```

```

alignment_scores:
  Quality: 50.00      Length: 10
  Ratio: 5.000       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
  US-08-653-294-29 x HS44032S1 ..
  Align seg 1/1 to: HS44032S1 from: 1 to: 270

11 ArgGluAsnLeuArgThrAlaLeuArgTyr 20
|||||
222 CGAGAGAACCTGCGCACCAGCGCTCCGCTAC 251

seq_name: gb_pr4:AF035648

seq_documentation_block:
  LOCUS AF035648 395 bp DNA PRI 05-JAN-1999
  DEFINITION Homo sapiens MHC class I antigen HLA-B gene (HLA-B*4402 variant
  allele), partial cds.
  ACCESSION AF035648
  VERSION AF035648.1 GI:4104422

```

**THIS PAGE BLANK (USPTO)**





```

FT mat_peptide      1..1143  

FT FT              /*tag= b  

FT                 319..1143  

FT FT              /*tag= c  

PN EP-398539-A.  

PD 22-NOV-1990.  

PF 01-MAY-1990; 304715.  

PR 17-MAY-1989; US-353124.  

PA (AMGE-) AMGEN INC.  

PI Zukowski MM, Stabinsky Y, Levitt M;  

DR WPI: 88-307568/43.  

DR P-PSDB: P80744.  

PT New subtilisin analogues -  

PT have aminoacid(s) present in calcium binding site replaced by  

PT negatively charged aminoacid(s)  

PP Example 2; Table 1, Pages 22-24; 60pp; English.  

PS Genomic DNA was isolated from cells of B. subtilis strain QB127 (trpc2  

CC leuA8 sacuH200). The coding region of the aprA gene was sequenced and  

CC the results of the sequence are given in n80745. The specific identity  

CC of the initial 5 codons of the leader region is attributable to the  

CC report of Stahl, et al., J. Bacteriol., 158, 411-418, (1984) and Wong,  

CC et al P.N.A.S., 81, 1184-1188 (1984). There exist codon sequence  

CC differences from Stahl, et al., at amino acid positions 84 and 85.  

CC Specifically, Stahl, et al., reports a codon GTT (coding for valine) at  

CC amino acid position 84 while the codon GTA (also coding for valine)  

CC appears in n80745. Stahl, et al., also reports a codon AGC (coding for  

CC serine) at amino acid position 85 as opposed to the codon GCG (coding for  

CC alanine) in n80745. The patent concerns a novel subtilisin analogue  

CC which has an amino acid sequence of a naturally occurring Bacillus  

CC subtilisin which has been modified by having: one or more of the amino  

CC acids present in a calcium binding site of the naturally occurring  

CC bacillus subtilisin replaced by a negatively charged amino acid , and  

CC one or more of any Asn-Gly sequence of the naturally occurring Bacillus  

CC subtilisin deleted or replaced by a different amino acid. Pref. it is  

CC an analogue of subtilisin Carlsberg, subtilisin DY, subtilisin BPN', an  

CC aprA subtilin of B. subtilis or subtilisin from B. mesentericus. The  

CC subtilisin analogues exhibit improved thermal and pH stability,  

CC increased specific activity and broad substrate specificity thereby  

CC increasing the detergency of detergent formulations contg. such  

CC analogues.  

SQ Sequence     1220 BP;   355 A;    281 C;    283 G;    301 T;
```

---

**alignment\_scores:**

Quality:	46.00	Length:	17
Ratio:	2.875	Gaps:	0
Percent Similarity:	94.118	Percent Identity:	52.941

**alignment\_block:**

```
US-08-653-294-29 x N80745/rev ..
Align seg 1/1 to reverse of: N80745 from: 1 to: 1220
```

**seq\_name:** N\_Geneseq\_36:N70890  
  
**seq\_documentation\_block:**  
ID N70890 standard; DNA; 1220 BP.  
AC N70890;  
DT 26-FEB-1991 (first entry)  
DE Thermostable Bacillus subtilisin encoded by aprA gene.  
KW aprA gene; subtilisin; thermostable enzyme; protease; surfactant; ds.  
OS Bacillus subtilis.  
FH Key Location/Qualifiers  
FT cds 1..1143  
FT FT /\*tag= a  
FT mat\_peptide 316..1143  
FT FT /\*tag= b  
FT FT /label= aprA gene  
FN WO8704461-A.  
PD 30-JUL-1987.  
PF 07-JAN-1987; U00027.  
PR 15-JAN-1986; US-819241.  
PA (AMGE-) AMGEN.  
PI Stabinsky Y, Zukowski M;  
DR WPI: 87-221262/31.  
DR P-PSDB: P70555.  
PT Thermally stable and pH stable subtilisin analogues - produced by  
PT deleting or replacing at least one of the asparagine-glycine  
PT sequences.  
PS Disclosure; Table 1; 56pp; English.  
CC Modified subtilisin may be used in detergent compositions, it has  
CC an improved pH and heat stability. See also p70556.  
SQ Sequence 1220 BP; 355 A; 283 C; 281 G; 301 T;

```

alignment_scores:
  Quality: 46.00      Length: 17
  Ratio: 2.875       Gaps: 0
  Percent Similarity: 94.118      Percent Identity: 52.941

alignment_block:
  US-08-653-294-29 x N70890/rev ..
  Align seg 1/1 to reverse of: N70890 from: 1 to: 1220

      3 LeuAlaileArgLeuAsnGluArgGluAsnLeuArgThrAlaLeuAr 19
      ::::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
      381 GTAGCCTTGAGTGAAGAGCGCGCTTTAATTGAGAAATGCCATAAG 332

      19 g 19
      |
      331 G 331

seq_name: N_Geneseq_36:Q03536

seq_documentation_block:
  ID Q03536 standard; DNA; 1499 BP.
  AC Q03536;
  DT 16-AUG-1990 (first entry)
  DE Subtilisin gene.
  KW Carboxyl hydrolase; subtilisin; neutral protease; ds.
  OS Bacillus subtilis.
  FH Key Location/Qualifiers
  FT Cds 137..1280
  FT /*tag= a
  PN EP-357157-A.
  PD 7-MAR-1990.
  PF 22-JUN-1984; 202584.
  PR 24-JUN-1983; US-507419.
  PR 29-MAY-1984; US-614612.
  PR 29-MAY-1984; US-614615.
  PR 29-MAY-1984; US-614491.
  PR 29-MAY-1984; US-614616.
  PR 29-MAY-1984; US-614617.
  PR (GETH) Genentech Inc.
  PA Bott RR, Estell DA, Ferrari E, Henner DJ, Wells JA;
  DR WPI; 90-068909/10.
  DR P-PSDB; R03737.
  PT Mutant prokaryotic carbonyl hydrolase enzymes -
  PT obtd. by site-directed oligo-nucleotide mutagenesis, used in
  PT food processing and cleaning industries.
  PS Claim 16; Fig 7; 39pp; English.
  CC Probe derived from subtilisin gene was used to isolate carbonyl
  CC hydrolase gene, mutant versions of which exhibit different oxidative
  CC stability and/or pH activity.
  SQ Sequence 1499 BP; 443 A; 342 C; 328 G; 386 T;

alignment_scores:
  Quality: 46.00      Length: 17
  Ratio: 2.875       Gaps: 0
  Percent Similarity: 94.118      Percent Identity: 52.941

alignment_block:
  US-08-653-294-29 x Q03536/rev ..
  Align seg 1/1 to reverse of: Q03536 from: 1 to: 1499

      3 LeuAlaileArgLeuAsnGluArgGluAsnLeuArgThrAlaLeuAr 19
      ::::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
      517 GTAGCCTTGAGTGAAGAGCGCGCTTTAATTGAGAAATGCCATAAG 468

      19 g 19
      |
      467 G 467

seq_name: N_Geneseq_36:N71241

```

```

seq_documentation_block:
  ID N71241 standard; DNA; 1500 BP.
  AC N71241;
  DT 02-MAY-1991 (first entry)
  DE Subtilisin gene from Bacillus subtilis.
  KW Subtilisin; extracellular protease; amylase production; ss.
  OS Bacillus subtilis.
  FH Key Location/Qualifiers
  FT signal_peptide 137..205
  FT /*tag= a
  FT mat_peptide 455..1279
  FT /*tag= b
  FT misc_rna 206..454
  FT /*tag= c
  FT /*label= pro sequence
  PN EP-246678-A.
  PD 25-NOV-1987.
  PF 01-JAN-1987; 200690.
  PR 24-JUN-1983; US-507419.
  PR 29-MAY-1984; US-614491.
  PR 29-MAY-1984; US-614616.
  PR 29-MAY-1984; US-614612.
  PR 29-MAY-1984; US-614617.
  PR 29-MAY-1984; US-614615.
  PR 01-JAN-1987; EP-200690.
  PR (GETH) GENENTECH INC.
  PI Bott RR, Ferrari E, Wells JA, Estell DA, Henner DJ;
  DR WPI; 87-328920/47.
  DR P-PSDB; P1060.
  PT Bacillus strains not excreting subtilisin or neutral protease -
  PT obtd. by recombinant DNA procedures, useful for enzyme prodn.
  PT esp. of hydrolase(s) such as amylase
  PS Example; Fig 7; 71pp; English.
  CC B.subtilis il68 chromosomal DNA was digested with EcoRI. A single
  CC 6kb fragment hybridised to a fragment from the C-terminus of the
  CC subtilisin structural gene in pS4 (see N71240). It was ligated to
  CC EcoRI-cut pBS42 and the ligation mixture used to transform E.coli
  CC ATCC 31446. Plasmid DNA was prepared from a pooled suspension of
  CC transformant colonies and used to transform a protease deficient
  CC strain of B.subtilis (BG84). Plasmid DNA from protease producing
  CC colonies was digested with EcoRI and examined by Southern blot
  CC analysis to isolate the 6kb fragment. A positive clone was identified
  CC containing a plasmid designated pS168.1. Three HincII fragments
  CC and a HincIII-EcoRI fragment were ligated into M13 vectors and
  CC sequenced to obtain the entire subtilisin sequence.
  SQ Sequence 1500 BP; 443 A; 343 C; 328 G; 386 T;

alignment_scores:
  Quality: 46.00      Length: 17
  Ratio: 2.875       Gaps: 0
  Percent Similarity: 94.118      Percent Identity: 52.941

alignment_block:
  US-08-653-294-29 x N71241/rev ..
  Align seg 1/1 to reverse of: N71241 from: 1 to: 1500

      3 LeuAlaileArgLeuAsnGluArgGluAsnLeuArgThrAlaLeuAr 19
      ::::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
      517 GTAGCCTTGAGTGAAGAGCGCGCTTTAATTGAGAAATGCCATAAG 468

      19 g 19
      |
      467 G 467

seq_name: N_Geneseq_36:N70058

seq_documentation_block:
  ID N70058 standard; DNA; 1500 BP.
  AC N70058;
  DT 20-JAN-1991 (first entry)
  DE Bacillus subtilis subtilisin gene.

```

```

KW Bacillus subtilis; subtilisin; enzyme; ss.
FH Key Location/Qualifiers
FT Cds 137..1280
FT /*tag= a
FT /product=subtilisin
PN EP-247647-A.
PD 02-DEC-1987.
PF 22-JUN-1984; 200689.
PR 24-JUN-1983; US-507419.
PR 29-MAY-1984; US-614491.
PA (GETH) Genentech Inc.
PI Bott RR,
PI Ferrari E,
PI Wells JA,
PI Estell DA,
PI Henner DJ,
DR WPI; 87-336315/48.
DR P-PSDB; P70053.
PT DNA mutagenesis proces - by introducing restriction enzyme sites,
PT digesting and introducing oligonucleotides capable of annealing at
PT the restriction enzyme sites.
PS Disclosure; Fig 7; 61pp; English.
CC This recombinant sequence may be used for the industrial production
CC of subtilisin. See also N70057, N70059 and P70052, P70054.
SQ Sequence 1500 BP; 443 A; 342 C; 330 G; 385 T;

alignment_scores:
  Quality: 46.00 Length: 17
  Ratio: 2.875 Gaps: 0
Percent Similarity: 94.118 Percent Identity: 52.941

alignment_block:
US-08-653-294-29 x N70058/rev ..
Align seg 1/1 to reverse of: N70058 from: 1 to: 1500

3 LeuAlaIleArgLeuAsnGluArgGluAsnLeuArgThrAlaLeuAr 19
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
517 GTAGCCTTGAGAGTGAGAGCGCGCTTTAATTTCGGAATGCCATAG 468

19 g 19
467 G 467

seq_name: N_Geneseq_36:Q90042
seq_documentation_block:
ID Q90042 standard; DNA; 1500 BP.
AC Q90042;
DT 10-JAN-1996 (first entry)
DE Bacillus subtilisin gene.
KW Bacillus amyloliquefaciens subtilisin; carbonyl hydrolase; protease;
KW fusion protein; preproprotein; maturation; ss.
KW autoproteolysis; maturation; ss.
OS Bacillus subtilis.
FH Key Location/Qualifiers
FT Cds 137..1282
FT /*tag= a
FT /product= subtilisin
FT /transl_except= 137..139 a.a.: fMet
FT signal_peptide 137..454
FT /*tag= b
FT mat_peptide 455..1279
FT /*tag= c
FT US5411873-A.
PN 02-MAY-1995.
PF 29-MAY-1984; 614612.
PR 29-MAY-1984; US-614612.
PR 01-APR-1986; US-846627.
PR 27-FEB-1990; US-488433.
PR 11-AUG-1992; US-928697.
PA (GENV) GENENCOR INC.

```

```

PI Adams RM, Power SD, Powers DB, Wells JA, Yansura DG;
DR WPI; 95-178127/23.
DR P-PSDB; R74224.
PT Recovery of recombinant subtilisin mutants from host cells - by
PT treatment with active subtilisin to cleave mutant from its
PT pro-sequence.
PS Disclosure; Fig 2; 32pp; English.
CC The nucleotide sequence of the Bacillus subtilis subtilisin gene. The
CC gene is used in a method to produce a carbonyl hydrolase (subtilisin)
CC e.g. the B.amyloliquefaciens subtilisin (Q90041) or other heterologous
CC protein (produced as a fusion protein) e.g. human growth hormone, such
CC that the desired protein is translated as a preproprotein which can be
CC transported across the cell membrane but is not released as an
CC enzymatically functional protein until the application of an external
CC protease or a protease encoded by the host cell e.g. B.subtilis
CC subtilisin or neutral protease (Q90043). The preproprotein sequence is
CC mutated so that it is incapable of autoproteolytic maturation. The
CC B.amyloliquefaciens sequence was mutated using the primers Q90044-5 and
CC Q****-*, specifically at S221N, D32N, A48R or contained a deletion of
CC 166 amino acids from the C-terminus of the protein.
SQ Sequence 1500 BP; 443 A; 343 C; 328 G; 386 T;

alignment_scores:
  Quality: 46.00 Length: 17
  Ratio: 2.875 Gaps: 0
Percent Similarity: 94.118 Percent Identity: 52.941

alignment_block:
US-08-653-294-29 x Q90042/rev ..
Align seg 1/1 to reverse of: Q90042 from: 1 to: 1500

3 LeuAlaIleArgLeuAsnGluArgGluAsnLeuArgThrAlaLeuAr 19
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
517 GTAGCCTTGAGAGTGAGAGCGCGCTTTAATTTCGGAATGCCATAG 468

19 g 19
467 G 467

seq_name: N_Geneseq_36:N60475
seq_documentation_block:
ID N60475 standard; DNA; 1524 BP.
AC N60475;
DT 24-AUG-1991 (first entry)
DE Sequence of the apr [Bsu] gene encoding subtilisin.
KW Alkaline serine protease; exoprotease; enzyme; proteolytic;
KW degradation; Bacillus host strains; ss.
OS Bacillus subtilisin.
FH Key Location/Qualifiers
FT signal_peptide 283..369
FT /*tag= a
FT /transl_except= (pos:283..285,aa:Met)
FT cds 370..600
FT /*tag= b
FT /label= pro sequence
FT mat_peptide 601..1428
FT /*tag= c
FT WO8601825-A.
PN 27-MAR-1986.
PF 20-SEP-1985; 905017.
PR 21-SEP-1984; US-652968.
PR 13-MAY-1985; US-733379.
PR (GENE-) GENEX CORP.
PA Farnestock SR, Fisher KE;
DR WPI; 86-094078/14.
DR P-PSDB; P60571.
PT Bacillus strains with reduced extra-cellular protease levels -
PT useful as hosts for secretion of heterologous polypeptide(s) and
PT proteins.
PS Example; Fig 2; 30pp; English.

```



KW Human leukocyte antigen; transgenic; germ cells; somatic cells;  
 PN expression: ss.  
 PD J04091731-A.  
 PF 25-MAR-1992. 207329.  
 PR 03-AUG-1990; JP-207329.  
 PA (OLYU ) OLYMPLUS OPTICAL CO.  
 DR WPI: 92-342893/42.  
 PT Transgenic non-human mammalian HLA-Bw 52 gene - useful for  
 PT analysis of expression of gene structure, and prodn. of  
 PT mouse model of human disease  
 PS Disclosure: Fig 1; 8pp; Japanese.  
 CC The sequence shows the exon 2 alpha-1-domain of the human leukocyte  
 CC antigen-Bw 52 gene. The complete gene may be introduced into non-  
 CC human mammals, pref. rat or mouse, or their ancestors at the primary  
 CC developmental biological step via transplacental into the zygote or  
 CC embryo to generate transgenic non-human mammals incorporating the  
 CC HLA-Bw 52 gene in both their germ cells and somatic cells. Transgenic  
 CC non-human mammals contg. HLA-Bw 52 are useful for the analysis of  
 CC expression of the gene, its structure, and prodn. of mouse models of  
 CC human disease. See also Q29166-72.  
 SQ Sequence 270 BP; 59 A; 88 C; 86 G; 37 T;

alignment\_scores:  
 Quality: 44.00 Length: 10  
 Ratio: 4.889 Gaps: 0  
 Percent Similarity: 90.000 Percent Identity: 90.000

## alignment\_block:

US-08-653-294-29 x Q29167 ..

Align seg 1/1 to: Q29167 from: 1 to: 270

11 ArgGluAsnLeuArgThrAlaLeuArgTyr 20

|||||

222 CGAGAGACCTGCGGATCGCGTCCGCTAC 251

seq\_name: N\_Geneseq\_36:T42608

## seq\_documentation\_block:

ID T42608 standard; DNA; 285 BP.

AC T42608;

DT 07-AUG-1997 (first entry)

DE Synthetic gene, Lqhiv encoding scorpion neurotoxin.

KW neurotoxin; pest control; insect; acarid; nematode; Lqhiv; Lqhvi;

KW venom; Leirus quinquistriatus hebraeus; scorpion; lepidptera;

KW baculovirus; arthropod; ds.

OS Synthetic.

FH Key Location/Qualifiers

FT cds

FT 70..267

FT /\*tag= a

FT /note= "no start codon given"

FT W09636221-A2.

PN 21-NOV-1996.

PD 30-APR-1996; U06076.

PR 08-MAY-1995; US-435040.

PA (REGC ) UNIV CALIFORNIA.

PI Hammock BD, Herrmann R, Moskowitz H;

DR WPI: 97-011744/01.

DR P-PSDB; W06339.

PT Control of insects, acarids and nematodes - using novel scorpion

PT toxins or a combination of 2 or more insect toxins

PS Claim 1; Fig 1; 35pp; English.

CC A novel method for controlling pests selected from insects, acarids, and

CC nematodes, comprises treating the pests or their loci with at least two

CC different insect toxins, the source of the toxins being at least one

CC recombinant microbe, the toxins having non-overlapping binding sites at

CC an insect cellular membrane channel. The present sequence is a synthetic

CC gene for Lqhiv toxin, derived from the venom of Leirus quinquistriatus

CC hebraeus (Scorpion). Lqhiv is the most potent lepidpterous toxin

CC isolated from scorpion venom to date, whilst Lqhvi (W06340) has weak

CC mammal toxicity. In particular the insect toxin genes are expressed

CC by baculovirus vectors, as the baculoviruses infect only arthropods,  
 CC therefore pose little or no risk to humans, plants or the environment.  
 SQ Sequence 285 BP; 57 A; 92 C; 70 G; 56 T;

## alignment\_scores:

Quality: 44.00 Length: 19

Ratio: 2.750 Gaps: 0

Percent Similarity: 84.211 Percent Identity: 52.632

## alignment\_block:

US-08-653-294-29 x T42608 ..

Align seg 1/1 to: T42608 from: 1 to: 285

1 TyrArgLeuAlaIleArgLeuAsnGluArgGluAsnLeuArgThrAl 17

|||||

66 CACCGCGTGGCGCGCGCTACATCGCGACGACGAAGTGGGTACA 115

|||||

17 aLeuArg 19

|||||

116 CCTGGCG 122

seq\_name: N\_Geneseq\_36:Q01834

## seq\_documentation\_block:

ID Q01834 standard; DNA; 1086 BP.

AC Q01834;

DT 19-MAR-1991 (first entry)

DE Sequence encoding HLA-B51 antigen.

KW Probe; HLA class I DNA; Immunogen; ss.

OS Homo sapiens.

PN EP354580-A.

PD 14-FEB-1990.

PR 10-AUG-1989.

PF 11-AUG-1988; JP-200758.

PA (OLYU) Olympus Optical Co., Ltd.

PI Kano K, Takiguchi;

DR WPI: 90-046289/07.

PT New DNA for class I human leukocyte antigens and derived probes and

PT transformed cells, useful for DNA typing, as immunogens etc.

PS Claim 1; Page 11; 23pp; English.

CC The HLA class I DNA can be used as a source of probes for use in DNA

CC typing. Transformed cells, which are useful as immunogens, can be

CC obtained by introducing these DNAs into eucaryotic cells.

SQ Sequence 1086 BP; 224 A; 334 C; 356 G; 172 T;

## alignment\_scores:

Quality: 44.00 Length: 10

Ratio: 4.889 Gaps: 0

Percent Similarity: 90.000 Percent Identity: 90.000

## alignment\_block:

US-08-653-294-29 x Q01834 ..

Align seg 1/1 to: Q01834 from: 1 to: 1086

11 ArgGluAsnLeuArgThrAlaLeuArgTyr 20

|||||

294 CGAGAGACCTGCGGATCGCGTCCGCTAC 323

seq\_name: N\_Geneseq\_36:Q01822

## seq\_documentation\_block:

ID Q01822 standard; DNA; 1086 BP.

AC Q01822;

DT 19-MAY-1991 (first entry)

DE Sequence encoding HLA-Bw52 antigen.

KW Probe; HLA class I DNA; Immunogen; ss.

OS Homo sapiens.

PN Key

FT Location/Qualifiers

FT cds

FT 1..1086

```

FT      /*tag= a
EP-354580-A.
PN      14-FEB-1990.
PD      10-AUG-1989.
PF      11-AUG-1988; JP-200758.
PR      (OLYU) Olympus Optical Co., Ltd.
PI      Kano K, Takiguchi.
PA      WPI; 90-046289/07.
DR      P-PSDB; R03142.
PT      New DNA for class 1 human leucocyte antigens and derived probes and
PT      transformed cells, useful for DNA typing, as immunogens etc.
PS      Claim 2; p11-12; 23pp; English.
PT      The HLA class I DNA can be used as a source of probes for use in DNA
CC      typing. Transformed cells, which are useful as immunogens, can be
CC      obtained by introducing these DNAs into eucaryotic cells.
SQ      Sequence 1086 BP; 223 A; 335 C; 358 G; 170 T;

alignment_scores:
Quality: 44.00 Length: 10
Ratio: 4.889 Gaps: 0
Percent Similarity: 90.000 Percent Identity: 90.000

alignment_block:
US-08-653-294-29 x Q01822 ..
Align seg 1/1 to: Q01822 from: 1 to: 1086

11 ArgGluAsnLeuArgThrAlaLeuArgTyr 20
|||||
294 CGAGAGAACTGGGATCGGCTCCGCTAC 323

```

**THIS PAGE BLANK (USPTO)**





```

DEFINITION EST247158 tomato ovary, TAMU Lycopersicon esculentum cDNA clone
ACCESSION CLE018L8, mRNA sequence.
VERSION AI488819.1 GI:4384190
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Potatoe; Lycopersicon.
REFERENCE 1 (bases 1 to 647)
AUTHORS Alcala,J., Vrebalov,J., White,R., Matern,A.L., Vision,T.,
Holt,I.E., Liang,F., Upton,J., Ronning,C.M., Craven,M.B.,
Fujii,C.Y., Bowman,C.L., Nierman,W., Fraser,C.M., Venter,J.C.,
Martin,G.B., Tanksley,S.D. and Giovannoni,J.
Generation of ESTs from tomato carpel tissue
Unpublished (1999)
COMMENT On May 18, 1998 this sequence version replaced gi:3138293.
Contact: David Frisch
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 4366
Fax: 864 656 4293
Email: dfrisch@CLEMSON.EDU.
FEATURES
Source
1..647
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="CLE018L8"
/clone_lib="tomato ovary, TAMU"
/tissue_type="carpel"
/dev_stage="5 days pre-anthesis to 5 days post-anthesis"
/lab_host="XL1-Blue MRF"
/notes="Vector: pBlueScript SK(-); Site_1: EcoRI; Site_2:
XhoI; cLED - Tomato Carpel EST Library. OligodT-primed and
directionally cloned cDNA in vector Lambda Zap II with 5',
and 3' ends located at the EcoRI and XhoI sites,
respectively."
BASE COUNT 198 a 116 c 148 g 185 t
ORIGIN
|||||
alignment_scores:
Quality: 52.00 Length: 20
Ratio: 3.714 Gaps: 0
Percent Similarity: 70.000 Percent Identity: 60.000
seq_documentation_block:
US-08-653-294-29 x AI488819/rev ..
Align seg 1/1 to reverse of: AI488819 from: 1 to: 647
1 TTTArgLeuAlaIleArgLeuAsnGluArgGluAsnLeuArgThrAl 17
|||||
384 TATGTCGCATACGATTCGACAGATTCGGAACATTCAGGACAG 335
17 aLeuArgTyr 20
|||||
334 GTTGAGGTAT 325
seq_name: gb_est1:D35316
seq_documentation_block:
LOCUS D35316 360 bp. mRNA EST 08-AUG-1994
DEFINITION CELK019CZF Yuji Kohara unpublished cDNA Caenorhabditis elegans CDNA
clone yk19g12 5', mRNA sequence.
ACCESSION D35316
VERSION D35316.1 GI:526832
KEYWORDS EST.
SOURCE Caenorhabditis elegans.
ORGANISM Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE 1 (bases 1 to 360)
AUTHORS Kohara,Y., Mitsuki,H., Nishigaki,A., Motohashi,T., Sugimoto,A. and
Tabara,H.
Toward an expression map of the C.elegans genome
Unpublished (1994)
COMMENT Contact: Yuji Kohara
Gene Library Lab
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.
FEATURES
Source
1..360
/organism="caenorhabditis elegans"
/strain="CB1489 him-8(e1489)"
/db_xref="taxon:6239"
/clone="yk19g12"
/clone_lib="yuji kohara unpublished cDNA"
/sex="hermaphrodite, male"
/tissue_type="whole animal"
/dev_stage="varied"
BASE COUNT 111 a 87 c 82 g 77 t 3 others
ORIGIN
|||||
alignment_scores:
Quality: 50.00 Length: 17
Ratio: 3.571 Gaps: 0
Percent Similarity: 82.353 Percent Identity: 64.706
seq_documentation_block:
US-08-653-294-29 x D35316 ..
Align seg 1/1 to: D35316 from: 1 to: 360
2 ArgLeuAlaIleArgLeuAsnGluArgGluAsnLeuArgThrAlaLe 18
|||||
61 AGATCCCGGTAAAGTTGAGGAAGAAGATCAATCTGAGGATTATTCT 110
18 u 18
|
111 G 111
seq_name: gb_est1:D35321
seq_documentation_block:
LOCUS D35321 360 bp. mRNA EST 08-AUG-1994
DEFINITION CELK019G6F Yuji Kohara unpublished cDNA Caenorhabditis elegans CDNA
clone yk19g6 5', mRNA sequence.
ACCESSION D35321
VERSION D35321.1 GI:526826
KEYWORDS EST.
SOURCE Caenorhabditis elegans.
ORGANISM Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE 1 (bases 1 to 360)
AUTHORS Kohara,Y., Mitsuki,H., Nishigaki,A., Motohashi,T., Sugimoto,A. and
Tabara,H.
Toward an expression map of the C.elegans genome
Unpublished (1994)
COMMENT Contact: Yuji Kohara
Gene Library Lab
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.
FEATURES
Source
1..360
/organism="caenorhabditis elegans"
/strain="CB1489 him-8(e1489)"
/db_xref="taxon:6239"
/clone="yk19g12"
/clone_lib="yuji kohara unpublished cDNA"
/sex="hermaphrodite, male"
/tissue_type="whole animal"
/dev_stage="varied"

```

```

source
1..360
/organism="Caenorhabditis elegans"
/strain="CB1489 him-8(e1489)"
/db_xref="taxon:6239"
/clone="yk1996"
/clone_lib="Yuji Kohara unpublished cDNA"
/sex="hermaphrodite, male"
/tissue_type="whole animal"
/dev_stage="varied"
BASE COUNT      116 a      85 c      79 g      80 t
ORIGIN

alignment_scores
Quality: 50.00      Length: 17
Ratio: 3.571      Gaps: 0
Percent Similarity: 82.353      Percent Identity: 64.706

alignment_block:
US-08-653-294-29 x D35321 ..
Align seg 1/1 to: D35321 from: 1 to: 360

2 ArgLeuAlaileArgLeuAsnGluArgGluAsnLeuArgThralale 18
||| |||:|||||:|||||:|||||:|||||:|||||:|||||
79 AGATCCGCGTTAAGGTTGACGGAAGAAGATCCAAATCTGAGGATTATTCT 128

18 u 18
129 G 129

seq_name: gb_est1:D35579

seq_documentation_block:
LOCUS D35579 360 bp mRNA EST 08-AUG-1994
DEFINITION CELK02269F Yuji Kohara unpublished cDNA Caenorhabditis elegans cDNA
ACCESSION D35579
VERSION D35579.1 GI:527030
KEYWORDS EST.
SOURCE Caenorhabditis elegans.
ORGANISM Caenorhabditis elegans.
Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditidae;
Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
Kohara, Y., Mitsuki, H., Nishigaki, A., Motohashi, T., Sugimoto, A. and
Tabara, H.
Toward an expression map of the C.elegans genome
Unpublished (1994)
Contact: Yuji Kohara
Gene Library Lab
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.
FEATURES
Source
1..360
/organism="Caenorhabditis elegans"
/strain="CB1489 him-8(e1489)"
/db_xref="taxon:6239"
/clone="yk229"
/clone_lib="Yuji Kohara unpublished cDNA"
/sex="hermaphrodite, male"
/tissue_type="whole animal"
/dev_stage="varied"
BASE COUNT      105 a      84 c      87 g      77 t      6 others
ORIGIN

alignment_scores
Quality: 50.00      Length: 17
Ratio: 3.571      Gaps: 0

```

```

Percent Similarity: 82.353      Percent Identity: 64.706

alignment_block:
US-08-653-294-29 x D35579 ..
Align seg 1/1 to: D35579 from: 1 to: 360

2 ArgLeuAlaileArgLeuAsnGluArgGluAsnLeuArgThralale 18
||| |||:|||||:|||||:|||||:|||||:|||||:|||||
24 AGATCCGCGTTAAGGTTGACGGAAGAAGATCCAAATCTGAGGATTATTCT 73

18 u 18
74 G 74

seq_name: gb_est17:C69731

seq_documentation_block:
LOCUS C69731 360 bp mRNA EST 23-SEP-1997
DEFINITION Yuji Kohara unpublished cDNA Caenorhabditis elegans cDNA
ACCESSION C69731
VERSION C69731.1 GI:2431087
KEYWORDS EST.
SOURCE Caenorhabditis elegans.
ORGANISM Caenorhabditis elegans.
Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditidae;
Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
Kohara, Y., Motohashi, T., Tabara, H., Watanabe, H., Sugimoto, A.,
Sano, M., Miyata, A. and Nishigaki, A.
Expression map of the C.elegans genome
Unpublished (1996)
On Sep 12, 1996 this sequence version replaced gi:1316439.
Contact: Yuji Kohara
Gene Library Lab
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.
FEATURES
Source
1..360
/organism="Caenorhabditis elegans"
/strain="CB1489 him-8(e1489)"
/db_xref="taxon:6239"
/clone="yk368h10"
/clone_lib="Yuji Kohara unpublished cDNA"
/sex="hermaphrodite, male"
/tissue_type="whole animal"
/dev_stage="varied"
BASE COUNT      109 a      90 c      85 g      76 t
ORIGIN

alignment_scores
Quality: 50.00      Length: 17
Ratio: 3.571      Gaps: 0
Percent Similarity: 82.353      Percent Identity: 64.706

alignment_block:
US-08-653-294-29 x C69731 ..
Align seg 1/1 to: C69731 from: 1 to: 360

2 ArgLeuAlaileArgLeuAsnGluArgGluAsnLeuArgThralale 18
||| |||:|||||:|||||:|||||:|||||:|||||:|||||
44 AGATCCGCGTTAAGGTTGACGGAAGAAGATCCAAATCTGAGGATTATTCT 93

18 u 18
94 G 94

```

```

seq_name: gb_est17:C69775
seq_documentation_block: 360 bp mRNA EST 29-SEP-1997
LOCUS C69775
DEFINITION C69775 Yuji Kohara unpublished cDNA Caenorhabditis elegans cDNA
clone yk366d11 5', mRNA sequence.
ACCESSION C69775
VERSION C69775.1 GI:2440300
KEYWORDS EST.
SOURCE Caenorhabditis elegans.
ORGANISM Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditidae;
Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE 1 (bases 1 to 360)
AUTHORS Kohara,Y., Motohashi,T., Tabara,H., Watanabe,H., Sugimoto,A.,
Sano,M., Miyata,A. and Nishigaki,A.
TITLE Expression map of the C.elegans genome
JOURNAL Unpublished (1996)
COMMENT Contact: Yuji Kohara
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.
Location/Qualifiers
FEATURES
source
1..360
/organism="Caenorhabditis elegans"
/strain="CB1489 him-8(e1489)"
/db_xref="taxon:6239"
/clone_lib="Yuji Kohara unpublished cDNA"
/sex="hermaphrodite, male"
/tissue_type="whole animal"
/dev_stage="varied"
BASE COUNT 112 a 88 c 85 g 75 t
ORIGIN
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

```

```

TITLE Expression map of the C.elegans genome
JOURNAL Unpublished (1996)
COMMENT On May 18, 1995 this sequence version replaced gi:810965.
Contact: Yuji Kohara
Gene Library Lab
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.
Location/Qualifiers
FEATURES
source
1..360
/organism="Caenorhabditis elegans"
/strain="CB1489 him-8(e1489)"
/db_xref="taxon:6239"
/clone_lib="Yuji Kohara unpublished cDNA"
/sex="hermaphrodite, male"
/tissue_type="whole animal"
/dev_stage="varied"
BASE COUNT 115 a 87 c 85 g 72 t 1 others
ORIGIN
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

```



Ratio:	3.692	Gaps:	0
Percent Similarity:	92.857	Percent Identity:	57.143

alignment\_block:

Align seq 1/1 to: A0655562 from: 1 to: 627

1 TyrArgLeuAlaIleArgLeuAsnGluArgArgGluAsnLeu 14  
 ||||| ::::::::::::::::::::|||  
 123 TATAGAGGTGTGTAGSATGTAATAAAAAAGGCGAATCTG 164

seq\_name: gb\_gss14:AQ576359

seq_documentation_block:	693 bp	DNA	GSS	02-JUN-1999
LOCUS	AQ576359			
DEFINITION	nxhb0089E07f	CUGI Rice BAC Library	Oryza sativa	genomic clone
	nxhb0089E07f.	genomic survey	sequence.	

ACCESSION AQ576359  
VERSION AQ576359.1 GI:4976844  
KEYWORDS GSS.

SOURCE	ORGANISM	ORGANISM
	<i>Oryza sativa</i>	<i>Oryza sativa</i>

REFERENCE  
AUTHORS  
1 (bases 1 to 693)  
Wing, R.A. and Dean, R.A.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Tel: 864 656 7288

**Fax: 864 656 4293**

Email: [rwing@clemson.edu](mailto:rwing@clemson.edu)

Seq primer: TAATACGACTCACTATAGGG

Class: BAC ends

High quality sequence

Location/Qualifiers  
1 693

```
1. :033
/organism="Oryza sativa"
```

```
1. .053
/organism="Oryza sativa"
/strain="Japonica"
/cultivar="Nipponbare"
/db_xref="taxon:4530"
/clone="nbx0089E07f"
```

```
/clone_lib="CUGI Rice B
```

```
/tissue_type="Leaf"
```

```
/lab_host="E. coli DH10B"
```

```
/note="Vector: pBelobAC11
```

Hindiii; Rice is one of t  
world half of the world

world. Half of the world  
inhabiting highly populat

and subtropics, rely on r

carbohydrate. Monocotyled

( $2n=24$ ) with a haploid g

(Arumuganathan and Earle,

genome of rice, three tim

Arabidopsis, makes it suitable for facilitating positional cloning.

order to facilitate post- and genome sequencing of

and genome sequencing of  
library from *Oryza sativa*

library contains 36,864 c

of 128.5 Kb providing 10

**The deep coverage allows**

sequence with a probability

filters, each containing

122 a	232 c	219 a	119
represent the whole libra			

2271 B 2227 C 2227 D 2271







GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run On: February 8, 2000, 04:05:43 ; Search time 133.56 Seconds  
(without alignments)  
3.547 Million cell updates/sec

Title: US-08-653-294-30

Perfect score: 100

Sequence: 1 YRLATRLNERENLTALRY 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 188963 seqs, 23686106 residues

Total number of hits satisfying chosen parameters: 188963

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : A\_Geneseq\_36.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	100	100.0	20	1 R2910	HLA-B*2702 CTL modu
2	100	100.0	20	1 W33793	Peptide B2702.84-7
3	94	94.0	20	1 R2909	HLA-B*2702 CTL modu
4	94	94.0	20	1 R2908	HLA-B*2702 CTL modu
5	94	94.0	20	1 W33791	Peptide B2702.84-7
6	94	94.0	20	1 W33792	Peptide B2702.84-7
7	88	88.0	20	1 R2907	HLA-B*2702 CTL modu
8	88	88.0	20	1 R35428	HLA-B*2702 84-75/7
9	88	88.0	20	1 W33778	Immunomodulating d
10	75	75.0	20	1 R35430	HLA-B*2702 84-75/7
11	50	50.0	20	1 R33095	HLA-B*2702 CTL modu
12	50	50.0	20	1 R35426	HLA-B*2702 75-84(T)
13	50	50.0	20	1 W33788	Peptide B2702.75-8
14	46.5	46.5	20	1 R2911	HLA-B*2702 CTL modu
15	46.5	46.5	20	1 W33779	Immunomodulating d
16	45	45.0	25	1 R48286	Peptide fragment o
17	45	45.0	25	1 R33093	HLA-B*2702 CTL modu
18	45	45.0	25	1 R35422	HLA-B*2702 CTL modu
19	45	45.0	184	1 Y06801	Peptide seq ID No:
20	45	45.0	362	1 R31142	Sequence of HLA-B*5
21	45	45.0	362	1 R31144	Sequence of HLA-B*5
22	45	45.0	362	1 R12463	HLA-B*53 exon. HLA
23	44	44.0	10	1 R41208	Peptide fragment o
24	44	44.0	10	1 R33062	HLA-B*2702 CTL modu
25	44	44.0	10	1 R35413	Alpha1-helix of HL
26	44	44.0	10	1 W07512	T-cell modulating
27	44	44.0	10	1 W47265	Immunomodulatory p
28	44	44.0	10	1 W47266	Immunomodulatory p
29	44	44.0	10	1 W33784	Peptide B2702.75-8
30	44	44.0	15	1 R2912	HLA-B*2702 CTL modu
31	44	44.0	15	1 W33795	Peptide B2702.70-8
32	44	44.0	20	1 R2913	HLA-B*7 CTL modulat
33	44	44.0	20	1 R35415	HLA-B*7.84-75-84 Pa
34	44	44.0	20	1 W33790	Peptide B7.84-75/7

35	44	44.0	20	1 W33797	Peptide B7.84-75/7
36	44	44.0	25	1 R41205	Peptide fragment o
37	44	44.0	25	1 R83090	HLA-B*2702 CTL modu
38	44	44.0	25	1 R35416	HLA-B*2702.60-84. C
39	44	44.0	25	1 W33794	Peptide B2702.60-8
40	43	43.0	10	1 W07522	T-cell modulating
41	43	43.0	1228	1 R77673	S-layer protein en
42	43	43.0	1228	1 W22862	Bacillus stearothe
43	42	42.0	580	1 W68402	FelV-A gag protein
44	42	42.0	1784	1 R05898	Gene product of fl
45	42	42.0	1784	1 R94427	FELV F6A provirus

ALIGNMENTS

RESULT 1

ID R92910 standard; peptide; 20 AA.  
AC R92910;  
DT 16-MAY-1996 (first entry)  
DE HLA-B\*2702 CTL modulating peptide (B2702.84-75(T)/75-84(T)).  
KW cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;  
KW immunosuppressant; graft versus host disorder; transplantation; therapy;  
KW class I MHC; HLA-B\*2702.  
OS Synthetic.  
PN W09526979-A1.  
PD 12-OCT-1995.  
PF 05-APR-1995; U04349.  
PR 05-APR-1994; US-222851.  
PA (STRD) UNIV LELAND STANFORD JUNIOR.  
PI Clayberger C, Krensky AM, Parham P;  
DR WPI: 95-358582/46.  
PT Extension of acceptance period of transplants from MHC unmatched  
PT donor hosts - using Class I B75-84 MHC antigen of the recipient  
PT host  
PS Example 15; Page 36; 80pp; English.  
CC R83061-R83085, R83090-R83096 and R92907-R92913 represent fragments of  
CC class I major histocompatibility complex (MHC) antigens. This sequence  
CC is an inverted dimer of residues 75-84 of the alpha-1 domain of the class  
CC I MHC HLA-B\*2702. These sequences can be used to extend the period of  
CC acceptance by a recipient of a transplant from an MHC unmatched donor.  
CC The peptides are administered to a patient in conjunction with a  
CC subtherapeutic amount of an immunosuppressant. This is administered to  
CC the patient for a limited period of time (compared to the lifetime  
CC administration for current treatments). The peptides particularly  
CC modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)  
CC of the patient.  
SQ Sequence 20 AA:

Query Match 100.0%; Score 100; DB 1; Length 20;  
Best Local Similarity 100.0%; Pred. No. 3.6e-10;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YRLATRLNERENLTALRY 20  
DB 1 YRLATRLNERENLTALRY 20

RESULT 2

ID W33793 standard; peptide; 20 AA.  
AC W33793;  
DT 19-JUN-1998 (first entry)  
DE Peptide B2702.84-75(T)/75-84r tested for immunomodulating activity.  
KW Immunomodulating dimer; immunosuppressant drug; CTL activation;  
KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;  
KW rejection.  
OS Synthetic.  
PN W09744351-A1.  
PD 27-NOV-1997.

PF 22-MAY-1997; U08689.  
 PR 24-MAY-1996; US-653294.  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI Beulow R, Clayberger C, Krensky AM;  
 DR WPI: 98-086530/08.  
 PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B  
 PT alpha-1 domain, used for preventing rejection of transplants or  
 PT treating autoimmune diseases  
 PS Example 1; Page 19; 41pp; English.  
 CC Peptides W3784-98 and W3778-9 were assayed for their immunomodulating  
 CC activity. A peptide-type compound or variant is claimed which has  
 CC immunomodulating activity, including the N-terminal acylated and/or  
 CC C-terminal amidated or esterified forms of up to 60 amino acids, where  
 CC the peptide-type compound comprises the formula: A-B, where A, B =  
 CC (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or V; aa77 =  
 CC D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a hydrophobic or  
 CC small amino acid; aa82 = R or L; aa83 = G or R; and aa represents amino  
 CC acid. The sequence in the brackets may optionally be absent or truncated  
 CC at any peptide type bond within the brackets. The compounds comprise  
 CC amino acid sequences related to a Class I HLA-B alpha domain (positions  
 CC 79-84). They can be used to inhibit cytotoxic T-lymphocytes (CTL) from  
 CC undesirably attacking cells in a host or in vitro. They can also be  
 CC used in combination with antigenic peptides or proteins of interest to  
 CC activate CTLs. They can also inhibit the proliferation of T cells in  
 CC response to anti-CD3. The peptide can be used for preventing rejection  
 CC of transplants or for treating autoimmune diseases, e.g. diabetes,  
 CC rheumatoid arthritis and lupus erythematosus. The products can also be  
 CC used for detection and diagnosis.  
 SQ Sequence 20 AA;

Query Match 100.0%; Score 100; DB 1; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 3.6e-10; Indels 0; Gaps 0;  
 Matches 20; Conservative 0; Mismatches 0;

QY 1 YRLATRLNERENLTALRY 20  
 |||||  
 DB 1 YRLATRLNERENLTALRY 20

## RESULT 3

R92909 ID R92909 standard; peptide; 20 AA.  
 AC R92909;  
 DT 16-MAY-1996 (first entry)  
 DE HLA-B2702 CTL modulating peptide (B2702.84-75/75-84(T)).  
 KW Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;  
 KW immunosuppressant; graft versus host disorder; transplantation; therapy;  
 KW Class I MHC; HLA-B2702.  
 OS Synthetic.  
 PN W09526979-A1.  
 PD 12-OCT-1995.  
 PF 05-APR-1995; U04349.  
 PR 03-APR-1994; US-222851.  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI Clayberger C, Krensky AM, Parham P;  
 DR WPI: 95-358582/46.  
 PT Extension of acceptance period of transplants from MHC unmatched  
 PT donor hosts - using Class I B75-84 MHC antigen of the recipient  
 PT host

PS Example 15; Page 36; 80pp; English.  
 CC R3061-R83085, R83090-R83096 and R92907-R92913 represent fragments of  
 CC Class I major histocompatibility complex (MHC) antigens. This sequence  
 CC is an inverted dimer of residues 75-84 of the alpha-1 domain of the class  
 CC I MHC HLA-B2702. These sequences can be used to extend the period of  
 CC acceptance by a recipient of a transplant from an MHC unmatched donor.  
 CC The peptides are administered to a patient in conjunction with a  
 CC subtherapeutic amount of an immunosuppressant. This is administered to  
 CC the patient for a limited period of time (compared to the lifetime  
 CC administration for current treatments). The peptides particularly  
 CC modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)  
 CC of the patient.  
 SQ Sequence 20 AA;

Query Match 94.0%; Score 94; DB 1; Length 20;  
 Best Local Similarity 95.0%; Pred. No. 3.4e-09;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 YRLATRLNERENLTALRY 20  
 |||||  
 DB 1 YRLATRLNERENLTALRY 20

## RESULT 4

R92908 ID R92908 standard; peptide; 20 AA.  
 AC R92908;  
 DT 16-MAY-1996 (first entry)  
 DE HLA-B2702 CTL modulating peptide (B2702.84-75(T)/75-84).  
 KW Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;  
 KW immunosuppressant; graft versus host disorder; transplantation; therapy;  
 KW Class I MHC; HLA-B2702.  
 OS Synthetic.  
 PN W09526979-A1.  
 PD 12-OCT-1995.  
 PF 05-APR-1995; U04349.  
 PR 05-APR-1994; US-222851.  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI Clayberger C, Krensky AM, Parham P;  
 DR WPI: 95-358582/46.  
 PT Extension of acceptance period of transplants from MHC unmatched  
 PT donor hosts - using Class I B75-84 MHC antigen of the recipient  
 PT host

Query Match 94.0%; Score 94; DB 1; Length 20;  
 Best Local Similarity 95.0%; Pred. No. 3.4e-09;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 YRLATRLNERENLTALRY 20  
 |||||  
 DB 1 YRLATRLNERENLTALRY 20

## RESULT 5

R92911 ID R92911 standard; peptide; 20 AA.  
 AC R92911;  
 DT 19-JUN-1998 (first entry)  
 DE Peptide B2702.84-75/75-84 tested for immunomodulating activity.  
 KW Immunomodulating dimer; immunosuppressant drug; CTL activation;  
 KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;  
 KW rejection.  
 OS Synthetic.  
 OS Homo sapiens.  
 PN W09744351-A1.  
 PD 27-NOV-1997.  
 PF 22-MAY-1997; U08689.  
 PR 24-MAY-1996; US-653294.  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI Beulow R, Clayberger C, Krensky AM;  
 DR WPI: 98-086530/08.

PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B  
 PT alpha-1 domain, used for preventing rejection of transplants or  
 PT treating autoimmune diseases  
 PS Example 1; Page 19; 41pp; English.  
 CC Peptides W33784-98 and W33778-9 were assayed for their immunomodulating  
 CC activity. A peptide-type compound or variant is claimed which has  
 CC immunomodulating activity, including the N-terminal acylated and/or  
 CC C-terminal amidated or esterified forms of up to 60 amino acids, where  
 CC the peptide-type compound comprises the formula: A-B, where A, B =  
 CC (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or V; aa77 =  
 CC D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a hydrophobic or  
 CC small amino acid; aa82 = R or L; aa83 = G or R; and aa represents amino  
 CC acid. The sequence in the brackets may optionally be absent or truncated  
 CC at any peptide type bond within the brackets. The compounds comprise  
 CC amino acid sequences related to a Class I HLA-B alpha1 domain (positions  
 CC 79-84). They can be used to inhibit cytotoxic T-lymphocytes (CTL) from  
 CC undesirably attacking cells in a host or in vitro. They can also be  
 CC used in combination with antigenic peptides or proteins of interest to  
 CC activate CTLs. They can also inhibit the proliferation of T cells in  
 CC response to anti-CD3. The peptide can be used for preventing rejection  
 CC of transplants or for treating autoimmune diseases, e.g. diabetes,  
 CC rheumatoid arthritis and lupus erythematosus. The products can also be  
 CC used for detection and diagnosis.  
 SQ Sequence 20 AA;

Query Match 94.0%; Score 94; DB 1; Length 20;  
 Best Local Similarity 95.0%; Pred. No. 3.4e-09;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 YRLATRLNERENRLTALRY 20  
 ||||| ||||| ||||| |||||  
 Db 1 YRLATRLNERENRLTALRY 20

RESULT 6  
 W33792 ID R92907 standard; peptide; 20 AA.  
 AC W33792;  
 DT 19-JUN-1998 (first entry)  
 DE Peptide B2702.84-75/75-84 tested for immunomodulating activity.  
 KW Immunomodulating dimer; immunosuppressant drug; CTL activation;  
 KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;  
 KW rejection.  
 OS Synthetic.  
 OS Homo sapiens.  
 PN W09744351-A1.  
 PD 27-NOV-1997.  
 PF 22-MAY-1997; U08689.  
 PR 24-MAY-1996; US-653294.  
 PA (STRD) UNIV LELAND STANFORD JUNIOR.  
 PI Beulow R, Clayberger C, Krensky AM;  
 DR WPI; 98-086530/08.  
 PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B  
 PT alpha-1 domain, used for preventing rejection of transplants or  
 PT treating autoimmune diseases  
 PS Example 1; Page 19; 41pp; English.  
 CC Peptides W33784-98 and W33778-9 were assayed for their immunomodulating  
 CC activity. A peptide-type compound or variant is claimed which has  
 CC immunomodulating activity, including the N-terminal acylated and/or  
 CC C-terminal amidated or esterified forms of up to 60 amino acids, where  
 CC the peptide-type compound comprises the formula: A-B, where A, B =  
 CC (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or V; aa77 =  
 CC D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a hydrophobic or  
 CC small amino acid; aa82 = R or L; aa83 = G or R; and aa represents amino  
 CC acid. The sequence in the brackets may optionally be absent or truncated  
 CC at any peptide type bond within the brackets. The compounds comprise  
 CC amino acid sequences related to a Class I HLA-B alpha1 domain (positions  
 CC 79-84). They can be used to inhibit cytotoxic T-lymphocytes (CTL) from  
 CC undesirably attacking cells in a host or in vitro. They can also be  
 CC used in combination with antigenic peptides or proteins of interest to  
 CC activate CTLs. They can also inhibit the proliferation of T cells in  
 CC response to anti-CD3. The peptide can be used for preventing rejection

CC of transplants or for treating autoimmune diseases, e.g. diabetes,  
 CC rheumatoid arthritis and lupus erythematosus. The products can also be  
 CC used for detection and diagnosis.  
 SQ Sequence 20 AA;

Query Match 94.0%; Score 94; DB 1; Length 20;  
 Best Local Similarity 95.0%; Pred. No. 3.4e-09;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 YRLATRLNERENRLTALRY 20  
 ||||| ||||| ||||| |||||  
 Db 1 YRLATRLNERENRLTALRY 20

RESULT 7  
 R92907 ID R92907 standard; peptide; 20 AA.  
 AC R92907;  
 DT 16-MAY-1996 (first entry)  
 DE HLA-B2702 CTL modulating peptide (B2702.84-75/75-84).  
 KW Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;  
 KW immunosuppressant; graft versus host disorder; transplantation; therapy;  
 KW Class I MHC; HLA-B2702.  
 OS Synthetic.  
 OS Homo sapiens.  
 PN W09526979-A1.  
 PD 12-OCT-1995.  
 PF 05-APR-1995; U04349.  
 PR 05-APR-1994; US-222851.  
 PA (STRD) UNIV LELAND STANFORD JUNIOR.  
 PI Clayberger C, Krensky AM, Parham P;  
 DR WPI; 95-358582/46.  
 PT Extension of acceptance period of transplants from MHC unmatched  
 PT donor hosts - using Class I B75-84 MHC antigen of the recipient  
 PT host  
 PS Example 15; Page 36; 80pp; English.  
 CC R83061-R83085, R83090-R83096 and R92907-R92914 represent fragments of  
 CC class I major histocompatibility complex (MHC) antigens. This sequence  
 CC is an inverted dimer of residues 75-84 of the alpha-1 domain of the class  
 CC I MHC HLA-B2702. These sequences can be used to extend the period of  
 CC acceptance by a recipient of a transplant from an MHC unmatched donor.  
 CC The peptides are administered to a patient in conjunction with a  
 CC subtherapeutic amount of an immunosuppressant. This is administered to  
 CC the patient for a limited period of time (compared to the lifetime  
 CC administration for current treatments). The peptides particularly  
 CC modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)  
 CC of the patient.  
 SQ Sequence 20 AA;

Query Match 88.0%; Score 88; DB 1; Length 20;  
 Best Local Similarity 90.0%; Pred. No. 3.2e-08;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 YRLATRLNERENRLTALRY 20  
 ||||| ||||| ||||| |||||  
 Db 1 YRLATRLNERENRLTALRY 20

RESULT 8  
 R95428 ID R95428 standard; peptide; 20 AA.  
 AC R95428;  
 DT 12-NOV-1996 (first entry)  
 DE HLA-B2702 84-75-84 palindromic.  
 KW HLA; p74; alpha1-helix; human-leucocyte-associated antigen; inhibitor;  
 KW T-cell lysate; membrane protein; mammal; heat shock protein; Hsc70; APC;  
 KW B cell; calcium influx; cytotoxic T lymphocyte; CTL; differentiation;  
 KW cytolysis; antigen presenting cell.  
 OS Synthetic.  
 PN W09513288-A1.  
 PD 18-MAY-1995.  
 PF 10-NOV-1994; U12985.

PR 10-NOV-1993; US-150493.  
 PA (STRD) UNIV LELAND STANFORD JUNIOR.  
 PI Clayberger C, Krensky AM;  
 DR WPI; 95-194027/25.  
 PT Compens. comprising lymphoid surface membrane proteins - which may  
 PT inhibit cytolytic activity and differentiation of CTLs.  
 PS Example; Page 12; 29pp; English.  
 CC R95413, and R95415-R95431 represent palindromes and fragments of  
 CC human-leucocyte-associated antigens. This sequence represents the  
 CC HLA-B2702 84-75-84 palindromic. These sequences can be used to isolate  
 CC the protein p74 from a T-cell lysate. p74 is a T-cell surface membrane  
 CC protein associated with T-cell activation in mammalian T-cells, and is  
 CC also immunologically cross reactive with the heat shock protein Hsc70.  
 CC p74 is found in a limited number of cell types, but is particularly  
 CC expressed on B and T cells. p74 can be isolated by lysis of a suitable  
 CC cell with an amphoteric detergent, and then passed through an affinity  
 CC column containing a covalently bound HLA-B2702 palindromic peptide.  
 CC Compositions comprising the extracellular fragment of p74 combined with  
 CC HLA-B2702.60-84 (see R95416), induces calcium influx, and inhibits  
 CC cytotoxic T lymphocyte (CTL) differentiation or cytotoxicity. Candidate  
 CC compounds can be screened for their effect on the cytolytic activity of  
 CC T-cells, by combining them with the extracellular portion of p74 and  
 CC determining the amount of binding between the candidate compound and p74.  
 CC Modulation of CTL activity can be inhibited in a cellular composition  
 CC containing T-cells and antigen presenting cells (APCs), by adding to the  
 CC mix the extracellular portion of p74, in an amount sufficient to compete  
 CC with p74 for the binding of the p74 ligand.  
 SQ Sequence 20 AA;

Query Match 88.0%; Score 88; DB 1; Length 20;  
 Best Local Similarity 90.0%; Pred. No. 3.2e-08;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 YRLATRLNERENLRALRY 20  
 DB 1 YRLAIRENLRERLRALRY 20

RESULT 9  
 W33778  
 ID W33778 standard; peptide; 20 AA.  
 AC W33778;  
 DE 19-JUN-1998 (first entry)  
 DT Immunomodulating dimer peptide #1.  
 KW Immunomodulating dimer; immunosuppressant drug; CTL activation;  
 KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;  
 OS Synthetic.  
 OS Homo sapiens.  
 PN W09744351-A1.  
 PD 27-NOV-1997.  
 PF 22-MAY-1997; U08689.  
 PR 24-MAY-1996; US-653294.  
 PA (STRD) UNIV LELAND STANFORD JUNIOR.  
 PI Beulow R, Clayberger C, Krensky AM;  
 DR WPI; 98-086530/08.  
 PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B  
 PT alpha-1 domain, used for preventing rejection of transplants or  
 PT treating autoimmune diseases  
 PT Claim 16; Page 35; 41pp; English.  
 CC This sequence represents a specifically claimed immunomodulating  
 CC dimer peptide of the invention. A peptide-type compound or variant is  
 CC claimed which has immunomodulating activity, including the N-terminal  
 CC acylated and/or C-terminal amidated or esterified forms of up to 60  
 CC amino acids, where the peptide-type compound comprises the formula; A-B,  
 CC where A, B = (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or  
 CC V; aa77 = D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a  
 CC hydrophobic or small amino acid; aa82 = R or L; aa83 = G or R; and aa  
 CC represents amino acid. The sequence in the brackets may optionally be  
 CC absent or truncated at any peptide type bond within the brackets. The  
 CC compounds comprise amino acid sequences related to a Class I HLA-B  
 CC alpha1 domain (positions 79-84). They can be used to inhibit cytotoxic

CC T-lymphocytes (CTL) from undesirably attacking cells in a host or in  
 CC vitro. They can also be used in combination with antigenic peptides or  
 CC proteins of interest to activate CTLs. They can also inhibit the  
 CC proliferation of T cells in response to anti-CD3. The peptide can be  
 CC used for preventing rejection of transplants or for treating autoimmune  
 CC diseases, e.g. diabetes, rheumatoid arthritis and lupus erythematosus.  
 CC The products can also be used for detection and diagnosis.  
 SQ Sequence 20 AA;

Query Match 88.0%; Score 88; DB 1; Length 20;  
 Best Local Similarity 90.0%; Pred. No. 3.2e-08;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 YRLATRLNERENLRALRY 20  
 DB 1 YRLAIRENLRERLRALRY 20

## RESULT 10

R95430  
 ID R95430 standard; peptide; 20 AA.  
 AC R95430;  
 DT 12-NOV-1996 (first entry)  
 DE HLA-B2702 84-75T/75-84T palindromic.  
 KW HLA; p74; alpha1-helix; human-leucocyte-associated antigen; inhibitor;  
 KW T-cell lysate; membrane protein; mammal; heat shock protein; Hsc70; APC;  
 KW B cell; calcium influx; cytotoxic T lymphocyte; CTL; differentiation;  
 KW cytotoxicity; antigen presenting cell.  
 OS Synthetic.  
 PN W09513288-A1.  
 PD 18-MAY-1995.  
 PF 10-NOV-1994; U12985.  
 PR 10-NOV-1993; US-150493.  
 PA (STRD) UNIV LELAND STANFORD JUNIOR.  
 PI Clayberger C, Krensky AM;  
 DR WPI; 95-194027/25.  
 PT Compens. comprising lymphoid surface membrane proteins - which may  
 PT inhibit cytolytic activity and differentiation of CTLs.  
 PS Example; Page 12; 29pp; English.  
 CC R95413, and R95415-R95431 represent palindromes and fragments of  
 CC human-leucocyte-associated antigens. This sequence represents the  
 CC HLA-B2702 84-75T/75-84T palindromic. These sequences can be used to  
 CC isolate the protein p74 from a T-cell lysate. p74 is a T-cell surface  
 CC membrane protein associated with T-cell activation in mammalian T-cells,  
 CC and is also immunologically cross reactive with the heat shock protein  
 CC Hsc70. p74 is found in a limited number of cell types, but is  
 CC particularly expressed on B and T cells. p74 can be isolated by lysis of  
 CC a suitable cell with an amphoteric detergent, and then passed through an  
 CC affinity column containing a covalently bound HLA-B2702 palindromic  
 CC peptide. Compositions comprising the extracellular fragment of p74  
 CC combined with HLA-B2702.60-84 (see R95416), induces calcium influx, and  
 CC inhibits cytotoxic T lymphocyte (CTL) differentiation or cytotoxicity.  
 CC Candidate compounds can be screened for their effect on the cytolytic  
 CC activity of T-cells, by combining them with the extracellular portion of  
 CC p74 and determining the amount of binding between the candidate compound  
 CC and p74. Modulation of CTL activity can be inhibited in a cellular  
 CC composition containing T-cells and antigen presenting cells (APCs), by  
 CC adding to the mix the extracellular portion of p74, in an amount  
 CC sufficient to compete with p74 for the binding of the p74 ligand.  
 SQ Sequence 20 AA;

Query Match 75.0%; Score 75; DB 1; Length 20;  
 Best Local Similarity 84.2%; Pred. No. 4.1e-06;  
 Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 YRLATRLNERENLRALRY 19  
 DB 1 YRLAIRENLRERLRALRY 19

## RESULT 11

R83095  
ID R83095 standard; peptide; 10 AA.  
AC R83095;  
DT 16-MAY-1996 (first entry)  
DE HLA-B\*2702 CTL modulating peptide (B2702.75-84(T)).  
KW Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;  
KW immunosuppressant; graft versus host disorder; transplantation; therapy;  
KW class I MHC; HLA-B\*2702.  
OS Synthetic.  
PN W09528979-A1.  
PD 12-OCT-1995.  
PF 05-APR-1995; U04349.  
PR 05-APR-1994; US-222851.  
PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
PI Clayberger C, Krensky AM, Parham P;  
DR WPI; 95-358582/46.  
PT Extension of acceptance period of transplants from MHC unmatched  
PT donor hosts - using Class I B\*5-84 MHC antigen of the recipient  
PT host  
PS Example 14; Page 34; 80pp; English.  
CC R83095, R83096 and R92907-R92913 represent fragments of  
CC class I major histocompatibility complex (MHC) antigens. This sequence  
CC corresponds to residues 75-84 of the alpha-1 domain of the class I MHC  
CC HLA-B\*2702. This sequence showed no inhibitory effect upon cytotoxic T  
CC lymphocytes (CTLs). These sequences can be used to extend the period of  
CC acceptance by a recipient of a transplant from an MHC unmatched donor.  
CC The peptides are administered to a patient in conjunction with  
CC a subtherapeutic amount of an immunosuppressant. This is administered to  
CC the patient for a limited period of time (compared to the lifetime  
CC administration for current treatments). The peptides particularly  
CC modulate (or inhibit) the activity of the CTLs of the patient.  
SQ Sequence 10 AA;

Query Match 50.0%; Score 50; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.021;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 RENLRTALRY 20  
DB 1 RENLRTALRY 10  
|||||

RESULT 12  
R95426  
ID R95426 standard; peptide; 10 AA.  
AC R95426;  
DT 12-NOV-1996 (first entry)  
DE HLA-B\*2702.75-84(T).  
KW HLA; p74; alpha1-helix; human-leucocyte-associated antigen; inhibitor;  
KW T-cell lysate; membrane protein; mammal; heat shock protein; Hsc70; APC;  
KW B cell; calcium influx; cytotoxic T lymphocyte; CTL; differentiation;  
KW Cytolysis; antigen presenting cell.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT misc\_difference 6 /note= "I6T mutation"  
PN W09513288-A1.  
PD 18-MAY-1995.  
PF 10-NOV-1994; U12985.  
PR 10-NOV-1993; US-150493.  
PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
PI Clayberger C, Krensky AM;  
DR WPI; 95-194027/25.  
PT Compns. comprising lymphoid surface membrane proteins - which may  
PT inhibit cytolytic activity and differentiation of CTLs.  
PS Example; Page 11; 29pp; English.  
CC R95413, and R95415-R95431 represent palindromes and fragments of  
CC human-leucocyte-associated antigens. This sequence represents the  
CC HLA-B\*2702.75-84(T). These sequences can be used to isolate the protein  
CC p74 from a T-cell lysate. p74 is a T-cell surface membrane protein  
CC associated with T-cell activation in mammalian T-cells, and is also  
CC immunologically cross reactive with the heat shock protein Hsc70. p74 is

found in a limited number of cell types, but is particularly expressed on  
B and T cells. p74 can be isolated by lysis of a suitable cell with an  
amphoteric detergent, and then passed through an affinity column  
containing a covalently bound HLA-B\*2702 palindromic peptide.  
Compositions comprising the extracellular fragment of p74 combined with  
HLA-B\*2702.60-84 (see R95416), induces calcium influx, and inhibits  
cytotoxic T lymphocyte (CTL) differentiation or cytotoxicity. Candidate  
compounds can be screened for their effect on the cytolytic activity of  
T-cells, by combining them with the extracellular portion of p74 and  
determining the amount of binding between the candidate compound and p74.  
Modulation of CTL activity can be inhibited in a cellular composition  
containing T-cells and antigen presenting cells (APCs), by adding to the  
mix the extracellular portion of p74, in an amount sufficient to compete  
with p74 for the binding of the p74 ligand.  
SQ Sequence 10 AA;

Query Match 50.0%; Score 50; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.021;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 RENLRTALRY 20  
DB 1 RENLRTALRY 10  
|||||

RESULT 13  
W33788  
ID W33788 standard; peptide; 10 AA.  
AC W33788;  
DT 19-JUN-1998 (first entry)  
DE Peptide B2702.75-84R80 tested for immunomodulating activity.  
KW Immunomodulating dimer; immunosuppressant drug; CTL activation;  
KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;  
KW rejection.  
OS Synthetic.  
OS Homo sapiens.  
PN W09744351-A1.  
PD 27-NOV-1997. U08689.  
PF 22-MAY-1997; US-653294.  
PR 24-MAY-1996; US-653294.  
PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
PI Beulow R, Clayberger C, Krensky AM;  
DR WPI; 98-086530/08.  
PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B  
PT alpha-1 domain, used for preventing rejection of transplants or  
PT treating autoimmune diseases  
PS Example 1; Page 19; 41pp; English.  
CC Peptides W33784-98 and W33778-9 were assayed for their immunomodulating  
CC activity. A peptide-type compound or variant is claimed which has  
CC immunomodulating activity, including the N-terminal acylated and/or  
CC C-terminal amidated or esterified forms of up to 60 amino acids, where  
CC the peptide-type compound comprises the formula: A-B, where A, B =  
CC (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or V; aa77 =  
CC D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a hydrophobic or  
CC small amino acid; aa82 = R or L; aa83 = G or R; and aa represents amino  
CC acid. The sequence in the brackets may optionally be absent or truncated  
CC at any peptide type bond within the brackets. The compounds comprise  
CC amino acid sequences related to a Class I HLA-B alpha domain (positions  
CC 79-84). They can be used to inhibit cytotoxic T-lymphocytes (CTL) from  
CC undesirably attacking cells in a host or in vitro. They can also be  
CC used in combination with antigenic peptides or proteins of interest to  
CC activate CTLs. They can also inhibit the proliferation of T cells in  
CC response to anti-CD3. The peptide can be used for preventing rejection  
CC of transplants or for treating autoimmune diseases, e.g. diabetes,  
CC rheumatoid arthritis and lupus erythematosus. The products can also be  
CC used for detection and diagnosis.  
SQ Sequence 10 AA;

Query Match 50.0%; Score 50; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.021;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RENTLALRY 20  
 Db 1 RENTLALRY 10

## RESULT 14

R92911 ID R92911 standard; peptide: 20 AA.

AC R92911; 16-MAY-1996 (first entry)

DE HLA-B2702 CTL modulating peptide (B2702.84-75/84-75).

KW Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;

KW immunosuppressant; graft versus host disorder; transplantation; therapy;

KW class I MHC; HLA-B2702.

OS Synthetic.

PN WO9526979-A1.

PD 12-OCT-1995.

PF 05-APR-1995; U04349.

PR 05-APR-1994; US-222851.

PA (STRD ) UNIV LELAND STANFORD JUNIOR.

PI Clayberger C, Krensky AM, Parham P;

DR WPI: 95-358582/46.

PT Extension of acceptance period of transplants from MHC unmatched

PT donor hosts - using Class I B75-84 MHC antigen of the recipient

PT host

PS Example 15; Page 36; 80pp; English.

CC R83061-R83085, R83090-R83096 and R92907-R92914 represent fragments of

CC class I major histocompatibility complex (MHC) antigens. This sequence

CC is a dimer of residues 75-84 of the alpha-1 domain of the class I MHC

CC HLA-B2702. These sequences can be used to extend the period of

CC acceptance by a recipient of a transplant from an MHC unmatched donor.

CC The peptides are administered to a patient in conjunction with a

CC subtherapeutic amount of an immunosuppressant. This is administered to

CC the patient for a limited period of time (compared to the lifetime

CC administration for current treatments). The peptides particularly

CC modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)

CC of the patient.

CC Sequence 20 AA;

Query Match 46.5%; Score 46.5; DB 1; Length 20;

Best Local Similarity 63.2%; Pred. No. 0.17;

Matches 12; Conservative 1; Mismatches 3; Indels 3; Gaps 1;

QY 1 YRLATRLNERENLTALR 19

Db 1 YRLAIRLNER---YRLAIR 16

## RESULT 15

W33779 ID W33779 standard; peptide: 20 AA.

AC W33779;

DT 19-JUN-1998 (first entry)

DE Immunomodulating dimer peptide #2.

KW Immunomodulating dimer; immunosuppressant drug; CTL activation;

KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;

KW rejection.

OS Synthetic.

OS Homo sapiens.

PN WO9744351-A1.

PD 27-NOV-1997.

PF 22-MAY-1997; U08689.

PR 24-MAY-1996; US-653294.

PA (STRD ) UNIV LELAND STANFORD JUNIOR.

PI Beulow R, Clayberger C, Krensky AM;

DR WPI: 98-086530/08.

PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B

PT alpha-1 domain, used for preventing rejection of transplants or

PT treating autoimmune diseases

PS Claim 16; Page 35; 41pp; English.

CC This sequence represents a specifically claimed immunomodulating

CC dimer peptide of the invention. A peptide-type compound or variant is  
 CC claimed which has immunomodulating activity, including the N-terminal  
 CC acylated and/or C-terminal amidated or esterified forms of up to 60  
 CC amino acids, where the peptide-type compound comprises the formula: A-B,  
 CC where A, B = (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or  
 CC V; aa77 = D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a  
 CC hydrophobic or small amino acid; aa82 = R or L; aa83 = G or R; and aa  
 CC represents amino acid. The sequence in the brackets may optionally be  
 CC absent or truncated at any peptide type bond within the brackets. The  
 CC compounds comprise amino acid sequences related to a Class I HLA-B  
 CC alpha1 domain (positions 79-84). They can be used to inhibit cytotoxic  
 CC T-lymphocytes (CTL) from undesirably attacking cells in a host or in  
 CC vitro. They can also be used in combination with antigenic peptides or  
 CC proteins of interest to activate CTLs. They can also inhibit the  
 CC proliferation of T cells in response to anti-CD3. The peptide can be  
 CC used for preventing rejection of transplants or for treating autoimmune  
 CC diseases, e.g. diabetes, rheumatoid arthritis and lupus erythematosus.  
 CC The products can also be used for detection and diagnosis.  
 CC Sequence 20 AA;

Query Match 46.5%; Score 46.5; DB 1; Length 20;

Best Local Similarity 63.2%; Pred. No. 0.17;

Matches 12; Conservative 1; Mismatches 3; Indels 3; Gaps 1;

QY 1 YRLATRLNERENLTALR 19

Db 1 YRLAIRLNER---YRLAIR 16

Search completed: February 8, 2000, 04:05:43

Job time: 9360 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: February 7, 2000, 18:04:39 ; Search time 111.22 Seconds  
(without alignments)  
8.482 Million cell updates/sec

Title: US-08-653-294-30  
Perfect score: 100  
Sequence: 1 YRLATRLNERENLTALRY 20

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 142080 seqs, 47169319 residues

Total number of hits satisfying chosen parameters: 142080

Minimum DB seq length: 0  
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database : PIR\_62:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	51	51.0	355	137516	HLA-B alpha-chain
2	51	51.0	362	S25415	class I histocompa
3	51	51.0	362	A45850	MHC class I histoc
4	51	51.0	362	I61861	MHC HLA-B*44.2 cha
5	51	51.0	362	I54442	MHC class I histoc
6	48	48.0	137	I80174	class I histocompa
7	46	46.0	359	HLHUB4	MHC class I histoc
8	46	46.0	362	I54457	MHC class I lympho
9	45	45.0	273	I38509	MHC class I histoc
10	45	45.0	274	I54463	MHC HLA-B*38 chain
11	45	45.0	362	B30345	MHC class I histoc
12	45	45.0	362	A45834	MHC class I histoc
13	45	45.0	362	I84486	transmembrane gly
14	45	45.0	362	A30345	MHC class I histoc
15	45	45.0	362	I59633	MHC HLA-B transmem
16	45	45.0	362	S24434	class I histocompa
17	45	45.0	362	I37120	MHC class I histoc
18	45	45.0	483	S25606	bleomycin hydrolas
19	44	44.0	292	I57806	MHC H-2K-kml mRNA
20	44	44.0	354	I59308	class I histocompa
21	44	44.0	354	I80168	class I histocompa
22	44	44.0	354	I80167	class I histocompa
23	44	44.0	355	I80169	class I histocompa
24	44	44.0	355	I80171	class I histocompa
25	44	44.0	359	I1HLH012	MHC class I histoc
26	44	44.0	362	HLHUB8	MHC class I histoc
27	44	44.0	362	B45876	class I histocompa
28	44	44.0	362	JH0341	class I histocompa
29	44	44.0	362	JH0539	class I histocompa
30	44	44.0	362	JH0540	class I histocompa

31 44 44.0 362 2 I62045 gene HLA B-1517 pr  
32 44 44.0 362 2 I84490 lymphocyte antigen  
33 44 44.0 362 2 I37521 HLA-B\*57.2 antigen  
34 44 44.0 363 2 S07113 class I histocompa  
35 44 44.0 383 2 S03537 class I histocompa  
36 44 44.0 384 2 D35997 MHC class I histoc  
37 44 44.0 365 2 S77963 MHC class I histoc  
38 44 44.0 365 2 I54416 HLA-AW24 protein -  
39 44 44.0 365 2 I54493 MHC class I histoc  
40 44 44.0 369 1 HLMSKK MHC class I histoc  
41 43 43.0 368 2 A60854 MHC class I histoc  
42 43 43.0 368 2 I49712 H-2K-s - mouse  
43 43 43.0 368 2 I49713 H-2K-sml - mouse  
44 43 43.0 1228 2 I40468 surface layer prot  
45 43 43.0 2077 1 WZBE24 240K tegument prot

## ALIGNMENTS

RESULT 1  
I37516  
HLA-B alpha-chain - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 04-Oct-1996 #sequence\_revision 04-Oct-1996 #text\_change 23-Jul-1999  
C:Accession: I37516  
R:Gauchat-Feiss, D.; Breur-Vriesendorp, B.S.; Rufer, N.; Jeannet, M.; Roosnek, E.; T1  
Tissue Antigens 44, 261-264, 1994  
A:Title: Sequencing of a novel functional HLA-B\*44 subtype differing in two residues 1  
A:Reference number: I37516; MUID:95176328  
A:Accession: I37516  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-355 <RES>  
A:Cross-references: EMBL:X75953; NID:g791007; PIDN:CRAA53566.1; PID:g791008  
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 51.0%; Score 51; DB 2; Length 355;  
Best Local Similarity 68.8%; Pred. No. 1-2;  
Matches 11; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 TRLNERENLTALRY 20  
| | | | | | | | | |  
Db 93 TQTQYRENLRTALRY 108

RESULT 2  
S25415  
class I histocompatibility antigen HLA-B\*4403 alpha chain - human  
C:Species: Homo sapiens (man)  
C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 23-Jul-1999  
C:Accession: S25415  
R:Fleischhauer, K.; Kernan, N.A.; Dupont, B.; Yang, S.Y.  
Tissue Antigens 37, 133-137, 1991  
A:Title: The two major subtypes of HLA-B\*44 differ for a single amino acid in codon 15  
A:Reference number: S25415; MUID:91335451  
A:Accession: S25415  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-362 <FLE>  
A:Cross-references: EMBL:X64366; NID:g32178; PIDN:CRAA45718.1; PID:g32179  
C:Genetics:  
A:Gene: GDB:HLA-B  
A:Cross-references: GDB:120048; OMIM:142830  
A:Map position: 6p21.3-6p21.3  
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology  
C:Keywords: transmembrane protein  
F:220-285/Domain: immunoglobulin homology <IMM>

Query Match 51.0%; Score 51; DB 2; Length 362;  
Best Local Similarity 68.8%; Pred. No. 1-2;





A:Cross-references: GB:M15470; NID:g187680; PIDN:AAA59619.1; PID:g386883  
C:Genetics:

A:Gene: GDB:HLA-B  
A:Cross-references: GDB:120048; OMIM:142830  
A:Map position: 6p21.3-6p21.3  
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology  
C:Keywords: duplication; glycoprotein; heterodimer; transmembrane protein; transplantati  
F:1-21/Domain: signal sequence (fragment) #status predicted <SIG>  
F:22-359/Product: class I histocompatibility antigen HLA-B44 alpha chain #status predict  
F:22-304/Domain: extracellular #status predicted <EXT>  
F:22-111/Domain: alpha-1 <EX1>  
F:112-203/Domain: alpha-2 <EX2>  
F:217-282/Domain: immunoglobulin homology <IMM>  
F:305-328/Domain: transmembrane #status predicted <TM>  
F:329-359/Domain: intracellular #status predicted <INT>  
F:107/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 46.0%; Score 46; DB 1; Length 359;  
Best Local Similarity 62.5%; Pred. No. 7.3;  
Matches 10; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 5 TRLNERENLRALRY 20  
| | | | | | | |  
Db 90 TNTQTYRENLRALRY 105

RESULT 8  
I54457

MHC class I lymphocyte antigen - human  
C:Species: Homo sapiens (man)  
C:Date: 07-Jun-1996 #sequence\_revision 07-Jun-1996 #text\_change 23-Jul-1999  
C:Accession: I54457  
R:Pohla, H.; Kuon, W.; Tabaczewski, P.; Doerner, C.; Weiss, E.H.  
Immunogenetics 29, 297-307, 1989  
A:Title: Allelic variation in HLA-B and HLA-C sequences and the evolution of the HLA-B a  
A:Reference number: I54457; MUID:89233295  
A:Accession: I54457  
A:Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: mRNA  
A:Residues: 1-362 <RES>  
A:Cross-references: GB:M28205; NID:g576470; PIDN:AAA57145.1; PID:g576471  
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 46.0%; Score 46; DB 2; Length 362;  
Best Local Similarity 62.5%; Pred. No. 7.3;  
Matches 10; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 5 TRLNERENLRALRY 20  
| | | | | | | |  
Db 93 TNTQTYRENLRALRY 108

RESULT 9  
I38509

MHC class I histocompatibility antigen - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 06-Sep-1996 #sequence\_revision 06-Sep-1996 #text\_change 23-Jul-1999  
C:Accession: I38509  
R:Cereb, N.; Choi, J.W.; Riu, K.Z.; Yang, S.Y.  
Tissue Antigens 44, 271-273, 1994  
A:Title: HLA-B\*5105, a newly identified B51 IEF variant.  
A:Reference number: I38509; MUID:95176331  
A:Accession: I38509  
A:Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: mRNA  
A:Residues: 1-273 <RES>  
A:Cross-references: EMBL:U06697; NID:g469544; PIDN:AAA92997.1; PID:g469545  
C:Genetics:  
A:Gene: GDB:HLA-B  
A:Cross-references: GDB:120048; OMIM:142830  
A:Map position: 6p21.3-6p21.3

C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 45.0%; Score 45; DB 2; Length 273;  
Best Local Similarity 62.5%; Pred. No. 7.8;  
Matches 10; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 5 TRLNERENLRALRY 20  
| | | | | | | |  
Db 68 TNTQTYRENLRALRY 83

RESULT 10  
I54463

MHC HLA-B38 chain - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 07-Jun-1996 #sequence\_revision 07-Jun-1996 #text\_change 23-Jul-1999  
C:Accession: I54463  
R:Muehler, C.A.; Engler-Blum, G.; Gekeler, V.; Steiert, I.; Weiss, E.; Schmidt, H.  
Immunogenetics 30, 200-207, 1989  
A:Title: Genetic and serological heterogeneity of the supertypic HLA-B locus specific  
A:Reference number: I54463; MUID:89379286  
A:Accession: I54463  
A:Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: DNA  
A:Residues: 1-274 <RES>  
A:Cross-references: GB:M29864; NID:g187674; PIDN:AAA36222.1; PID:g187675  
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 45.0%; Score 45; DB 2; Length 274;  
Best Local Similarity 62.5%; Pred. No. 7.8;  
Matches 10; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 5 TRLNERENLRALRY 20  
| | | | | | | |  
Db 69 TNTQTYRENLRALRY 84

RESULT 11  
B30345

MHC class I histocompatibility antigen HLA-Bw52 precursor - human  
C:Species: Homo sapiens (man)  
C:Date: 29-Jan-1990 #sequence\_revision 29-Jan-1990 #text\_change 16-Feb-1997  
C:Accession: B30345  
R:Hayashi, H.; Ennis, P.D.; Ariga, H.; Salter, R.D.; Parham, P.; Kano, K.; Takiguchi, J.  
Immunol. 142, 306-311, 1989  
A:Title: HLA-B51 and HLA-Bw52 differ by only two amino acids which are in the helical  
A:Reference number: A30345; MUID:89080265  
A:Accession: B30345  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-362 <HAY>  
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology  
C:Keywords: transmembrane protein  
F:220-285/Domain: immunoglobulin homology <IMM>

Query Match 45.0%; Score 45; DB 2; Length 362;  
Best Local Similarity 62.5%; Pred. No. 11;  
Matches 10; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 5 TRLNERENLRALRY 20  
| | | | | | | |  
Db 93 TNTQTYRENLRALRY 108

RESULT 12  
A45834

MHC class I histocompatibility antigen HLA-B53 alpha chain precursor - human  
C:Species: Homo sapiens (man)  
C:Date: 03-Jun-1993 #sequence\_revision 03-Jun-1993 #text\_change 23-Jul-1999  
C:Accession: A45834

R:Hayashi, H.; Ooba, T.; Nakayama, S.; Sekimata, M.; Kano, K.; Takiguchi, M.  
Immunogenetics 32, 195-199, 1990  
A:Title: Allelopecificities between HLA-B\*53 and HLA-B\*35 are generated by substitution of  
A:Reference number: A45834; MUID:91033941  
A:Accession: A45834  
A:Molecule type: DNA  
A:Residues: 1-362 <HAY>  
A:Cross-references: GB:M58636; NID:q187756; PIDN:AAA36228.1; PID:q187757; GB:M33574  
A:Note: this allele is designated B\*5301  
C:Genetics:  
A:Gene: GDB:HLA-B  
A:Cross-references: GDB:120048; OMIM:142830  
A:Map position: 6p21.3-6p21.3  
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology  
C:Keywords: glycoprotein; heterodimer; transmembrane protein  
F:1-24/Domain: signal sequence #status predicted <Sig>  
F:220-285/Domain: immunoglobulin homology <IMM>  
F:110/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 45.0%; Score 45; DB 2; Length 362;  
Best Local Similarity 62.5%; Pred. NO. 11;  
Matches 10; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 5 TRLNERENLRALRY 20  
| | | | | | | | | |  
Db 93 TQTQYRENLRALRY 108

RESULT 13  
184486  
transmembrane glycoprotein - human  
C:Species: Homo sapiens (man)  
C:Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 23-Jul-1999  
C:Accession: 184486  
R:Hildebrand, W.H.; Domena, J.D.; Shen, S.Y.; Lau, M.; Terasaki, P.I.; Bunce, M.; Marsh,  
Tissue Antigens 43, 209-218, 1994  
A:Title: HLA-B\*51: a widespread and diverse family of HLA-B alleles.  
A:Reference number: 138421; MUID:94367483  
A:Accession: 184486  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-362 <RES>  
A:Cross-references: GB:L15005; NID:g493154; PIDN:AAA56832.1; PID:g493155  
C:Genetics:  
A:Gene: HLA-B\*1513  
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology  
C:Keywords: glycoprotein

Query Match 45.0%; Score 45; DB 2; Length 362;  
Best Local Similarity 62.5%; Pred. NO. 11;  
Matches 10; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 5 TRLNERENLRALRY 20  
| | | | | | | | | |  
Db 93 TQTQYRENLRALRY 108

RESULT 14  
A30345  
MHC class I histocompatibility antigen HLA-B\*51 precursor - human  
C:Species: Homo sapiens (man)  
C:Date: 29-Jan-1990 #sequence\_revision 29-Jan-1990 #text\_change 23-Jul-1999  
C:Accession: A30345; 168746; 137499  
R:Hayashi, H.; Ennis, P.D.; Ariga, H.; Salter, R.D.; Parham, P.; Kano, K.; Takiguchi, M.  
J. Immunol. 142, 306-311, 1989  
A:Title: HLA-B\*51 and HLA-B\*52 differ by only two amino acids which are in the helical re  
A:Reference number: A30345; MUID:89080265  
A:Accession: A30345  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-362 <HAY>

R:Pohla, H.; Kuon, W.; Tabaczewski, P.; Doerner, C.; Weiss, E.H.  
Immunogenetics 29, 297-307, 1989  
A:Title: Allelic variation in HLA-B and HLA-C sequences and the evolution of the HLA-  
A:Reference number: 134457; MUID:89233295  
A:Accession: 168746  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-362 <RES>  
A:Cross-references: GB:L41087; NID:g735900; PIDN:AAA64513.1; PID:g735902  
R:Steinle, A.; Schendel, D.J.  
Tissue Antigens 44, 268-270, 1994  
A:Title: HLA class I alleles of ICL 721 and 174 x CEM.T2 (T2).  
A:Reference number: 137499; MUID:95176330  
A:Accession: 137499  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-206 <RE2>  
A:Cross-references: EMBL:z46808; NID:g599783; PIDN:CAA86838.1; PID:g599784  
C:Genetics:  
A:Gene: GDB:HLA-B  
A:Cross-references: GDB:120048; OMIM:142830  
A:Map position: 6p21.3-6p21.3  
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology  
C:Keywords: transmembrane protein  
F:220-285/Domain: immunoglobulin homology <IMM>

Query Match 45.0%; Score 45; DB 2; Length 362;  
Best Local Similarity 62.5%; Pred. NO. 11;  
Matches 10; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 5 TRLNERENLRALRY 20  
| | | | | | | | | |  
Db 93 TQTQYRENLRALRY 108

RESULT 15  
159633  
MHC HLA-B transmembrane glycoprotein - human  
C:Species: Homo sapiens (man)  
C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 23-Jul-1999  
C:Accession: 159633  
R:Hildebrand, W.H.; Domena, J.D.; Parham, P.  
Tissue Antigens 41, 190-195, 1993  
A:Title: Primary structure shows HLA-B\*59 to be a hybrid of HLA-B\*55 and HLA-B\*51, and n  
A:Reference number: 159633; MUID:93369833  
A:Accession: 159633  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-362 <RES>  
A:Cross-references: GB:L07743; NID:g388314; PIDN:AAA59621.1; PID:g388315  
C:Genetics:  
A:Gene: GDB:HLA-B  
A:Cross-references: GDB:120048; OMIM:142830  
A:Map position: 6p21.3-6p21.3  
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology  
C:Keywords: glycoprotein

Query Match 45.0%; Score 45; DB 2; Length 362;  
Best Local Similarity 62.5%; Pred. NO. 11;  
Matches 10; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 5 TRLNERENLRALRY 20  
| | | | | | | | | |  
Db 93 TQTQYRENLRALRY 108

Search completed: February 7, 2000, 18:04:39  
Job time: 22205 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2000, 01:26:00 ; Search time 68.63 Seconds  
(without alignments)  
8.703 Million cell updates/sec

Title: US-08-653-294-30

Perfect score: 100

Sequence: 1 YRLATRLNERENLTALRY 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 82229 seqs, 29864866 residues

Total number of hits satisfying chosen parameters: 82229

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : SwissProt\_38.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	51	51.0	362	1	P30461
2	51	51.0	362	1	P30481
3	51	51.0	362	1	P30482
4	46	46.0	359	1	P10320
5	45	45.0	362	1	P30487
6	45	45.0	362	1	P18464
7	45	45.0	362	1	P30489
8	45	45.0	362	1	P30490
9	45	45.0	362	1	P30491
10	45	45.0	454	1	Q01532
11	44	44.0	359	1	P13750
12	44	44.0	362	1	P30379
13	44	44.0	362	1	P30380
14	44	44.0	362	1	P30381
15	44	44.0	362	1	P18465
16	44	44.0	362	1	P30497
17	44	44.0	362	1	P30499
18	44	44.0	362	1	P10193
19	44	44.0	362	1	P03447
20	44	44.0	365	1	P05534
21	44	44.0	365	1	P04223
22	44	44.0	369	1	P35825
23	43	43.0	1228	1	P30002
24	43	43.0	2077	1	P52340
25	43	43.0	2077	1	P12407
26	42	42.0	187	1	Q05631
27	42	42.0	351	1	P04322
28	42	42.0	414	1	P03337
29	42	42.0	425	1	P03340
30	42	42.0	536	1	P10262
31	42	42.0	580	1	P40850
32	41.5	41.5	630	1	P12044
33	41	41.0	162	1	Q42364
34	41	41.0	281	1	Q42364

## ALIGNMENTS

RESULT	ID	1B05_HUMAN	STANDARD;	PRT;	362 AA.
1	AC	P30461;			
DT	01-APR-1993	(Rel. 25, Created)			
DT	01-APR-1993	(Rel. 25, Last sequence update)			
DT	15-JUL-1998	(Rel. 36, Last annotation update)			
DE	HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-13 B*1301 ALPHA CHAIN				
DE	PRECURSOR (B13.1).				
GN	HLA-B OR HLAB.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;				
OC	Eutheria; Primates; Catarrhini; Hominidae; Homo.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE; 89235215.				
RA	PARHAM P., LAWOR D.A., LOMEN C.E., ENNIS P.D.;				
RT	"Diversity and diversification of HLA-A,B,C alleles.";				
RL	J. Immunol. 142:3937-3950(1989).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE; 88152906.				
RA	ZEMMOUR J., ENNIS P.D., PARHAM P., DUPONT B.;				
RT	"Comparison of the structure of HLA-B*47 to HLA-B13 and its relationship to 21-hydroxylase deficiency.";				
RL	Immunogenetics 27:281-287(1988).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE-BLOOD;				
RX	MEDLINE; 96053518.				
RA	LIN L., TOKUNAGA K., NAKAJIMA F., ISHIKAWA Y., KASHIWAKE K.,				
RA	TANAKA H., KUWATA S., SDELTSEVA E., AKAZA T., TADOKORO K.,				
RA	SHIBATA Y., CHANDANAYONG D., JUJI T.;				
RT	"Both HLA-B*1301 and B*1302 exist in Asian populations and are associated with different haplotypes.";				
RL	Hum. Immunol. 43:51-56(1995).				
CC	- - FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE IMMUNE SYSTEM.				
CC	- - SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-MICROGLOBULIN).				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).				
CC	-----				
CC	EMBL; M24041; AAA59660.1; -				
DR	EMBL; M19757; AAA52657.1; -				
DR	EMBL; P50291; BAA08822.1; -				
DR	HSSP; P30491; 1A1M.				
DR	MIM; 142830; -				
DR	PROSITE; PS00290; IG_MHC; 1.				
DR	PFAM; PF00047; Ig; 1.				
DR	PFAM; PF00129; MHC_I; 1.				

P18463 homo sapien  
P30485 homo sapien  
O59330 phanerochaete  
P49454 homo sapien  
P50084 saccharomyc  
O58377 methanococc  
P30467 homo sapien  
P03989 homo sapien  
P19373 homo sapien  
P10318 homo sapien  
O08136 homo sapien

35 41 41.0 362 1 1B29\_HUMAN  
36 41 41.0 362 1 1B45\_HUMAN  
37 41 41.0 380 1 LEU3\_PHACH  
38 41 41.0 3210 1 CENF\_HUMAN  
39 41 41.0 137 1 YG4V\_YEAST  
40 40 40.0 205 1 Y967\_METJA  
41 40 40.0 338 1 1B20\_HUMAN  
42 40 40.0 361 1 1B14\_HUMAN  
43 40 40.0 362 1 1B16\_HUMAN  
44 40 40.0 362 1 1B18\_HUMAN  
45 40 40.0 362 1 1B19\_HUMAN

```

FT CARBOHYD      110      110      BY SIMILARITY.
FT DISULFID      125      188      BY SIMILARITY.
FT DISULFID      227      283      BY SIMILARITY.
SQ SEQUENCE      362 AA; 40481 MW; D0AE6DD5 CRC32;

Query Match      51.0%; Score 51; DB 1; Length 362;
Best Local Similarity 58.8%; Pred. NO. 0.47;
Matches 11; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      5      TRLNERENLRALRY 20
      |      |||||
Db      93      TNTQTYENLRALRY 108

```

NO.	NAME	STANDARD;	PRT;
1B42	HUMAN		
ID	1B42_HUMAN		
AC	P30482		
		362	AA.

01-APR-1993 (Rel. 25, Last sequence update)  
01-APR-1993 (Rel. 25, Last sequence update)  
15-JUL-1998 (Rel. 36, Last annotation update)  
HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, BW-44(B-12) B\*4403 ALPHA CHAIN  
PRECURSOR.  
HLA-B OR HLAB.  
Homo sapiens (Human).  
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;  
Eutheria; Primates; Catarrhini; Hominidae; Homo.  
[1]  
SEQUENCE FROM N.A.  
MEDLINE; 91335451.  
FLEISCHHAUER K., KERNAN N.A., DUPONT B., YANG S.Y.;  
"The two major subtypes of HLA-B44 differ for a single amino acid in  
codon 156."  
Tissue Antigens 37:133-137(1991).

[2]  
RN SEQUENCE FROM N.A.  
RP MEDLINE; 96435470.  
RX ADAMS E.J., LITTLE A.M., ARNETT K.L., MCAULEY J.E., WILLIAMS R.C.,  
RA PARHAM P.;  
RT "Three new HLA-B alleles found in Mexican-Americans.";  
RL Tissue Antigens 46:414-416(1995).  
CC -I- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO  
CC THE IMMUNE SYSTEM.  
CC -I- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
CC MICROGLOBULIN).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL Outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: X64366; CAA45718.1; -

DR	EMBL; L42282; AAB51454.1;	-
DR	EMBL; L42283; AAB51455.1;	-
DR	PIR; S25415; S25415.	
DR	HSSP; P30491; 1ALM.	
DR	MIM; 142830;	-
DR	PROSITE; PS00290; IG_MHC; 1.	
DR	PFAM; PF00047; Ig; 1.	
DR	PFAM; PF00129; MHC I; 1.	
KW	MHC I; Transmembrane; Glycoprotein; Signal.	
FT	SIGNAL	1 <sup>24</sup>
FT	CHAIN	25 362
FT	DOMAIN	25 114
FT	DOMAIN	115 205
FT	DOMAIN	207 298
FT	DOMAIN	299 309
FT	TRANSMEM	310 333

FT DOMAIN 334 362 CYTOPLASMIC TAIL.  
 FT CARBOHYD 110 110 BY SIMILARITY.  
 FT DISULFID 125 188 BY SIMILARITY.  
 FT DISULFID 227 283 BY SIMILARITY.  
 SQ SEQUENCE 362 AA; 40479 MW; E7ZCE669 CRC32;

Query Match 51.0%; Score 51; DB 1; Length 362;

Best Local Similarity 68.8%; Pred. No. 0.47; Length 362;  
 Matches 11; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Oy 5 TRLNERENLTALRY 20  
 Db 93 TNTQTYRENLTALRY 108

RESULT 4  
 ID 1B40\_HUMAN STANDARD; PRT; 359 AA.  
 AC P10320;  
 DT 01-MAR-1989 (Rel. 10, Created)  
 DT 01-MAR-1989 (Rel. 10, Last sequence update)  
 DT 15-JUL-1999 (Rel. 36, Last annotation update)  
 DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, BW-44(B-12) B\*4401 ALPHA CHAIN  
 DE PRECURSOR (B44.1) (FRAGMENT).  
 GN HLA-B OR HLAB.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 86249389.  
 RA KOTTMANN A.H., SEEMANN G.H.A., GUESOW H.D., ROOS M.H.;  
 RT "DNA sequence of the coding region of the HLA-B44 gene.";  
 RL Immunogenetics 23:396-400(1986).

CC -1- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO  
 CC THE IMMUNE SYSTEM.  
 CC -1- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
 CC MICROGLOBULIN).  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; M15470; AAA59619.1;  
 CC PTR; A25295; HLHUB4.  
 CC HSP; P30491; 1A1M.  
 CC MIM; 142830;  
 CC PROSITE; PS00290; IG\_MHC; 1.  
 CC PFAM; PF00047; ig; 1.  
 CC PFAM; PF00129; MHC.I; 1.  
 CC MHC.I; Transmembrane; Glycoprotein; Signal.  
 CC NON\_TER 1 1  
 CC SIGNAL <1 21  
 CC CHAIN 22 359

FT DOMAIN 22 111 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,  
 FT DOMAIN 112 203 BW-44(B-12) B\*4401 ALPHA CHAIN.  
 FT DOMAIN 204 295 EXTRACELLULAR ALPHA-1.  
 FT DOMAIN 296 305 EXTRACELLULAR ALPHA-2.  
 FT TRANSMEM 306 329 EXTRACELLULAR ALPHA-3.  
 FT DOMAIN 330 359 CONNECTING PEPTIDE.  
 FT CARBOHYD 107 107 CYTOPLASMIC.  
 FT DISULFID 122 185 BY SIMILARITY.  
 FT DISULFID 224 280 BY SIMILARITY.  
 SQ SEQUENCE 359 AA; 40040 MW; 7C0DB32E CRC32;

Query Match

46.0%; Score 46; DB 1; Length 359;

Best Local Similarity 62.5%; Pred. No. 3;  
 Matches 10; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Oy 5 TRLNERENLTALRY 20  
 Db 90 TNTQTYRENLTALRY 105

RESULT 5  
 ID 1B47\_HUMAN STANDARD; PRT; 362 AA.  
 AC P30487;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 01-FEB-1996 (Rel. 33, Last annotation update)  
 DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B\*49(B-21) B\*4901 ALPHA CHAIN  
 DE PRECURSOR.  
 GN HLA-B OR HLAB.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 89235215.  
 RA PARHAM P., LAWLER D.A., LOMEN C.E., ENNIS P.D.;  
 RT "Diversity and diversification of HLA-A,B,C alleles.";  
 RL J. Immunol. 142:3937-3950(1989).  
 RN [2]  
 RP REVISION TO 78.  
 RX MEDLINE; 93056529.  
 RA HILDEBRAND W.H., MADRIGAL J.A., BELICH M.P., ZEMMOUR J., WARD F.E.,  
 RA WILLIAMS R.C., PARHAM P.;  
 RT "Serologic cross-reactivities poorly reflect allelic relationships in  
 RT the HLA-B\*2 and HLA-B\*21 groups. Dominant epitopes of the alpha 2  
 RT helix.";  
 RL J. Immunol. 149:3563-3568(1992).

CC -1- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO  
 CC THE IMMUNE SYSTEM.  
 CC -1- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
 CC MICROGLOBULIN).  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; M24037; AAA02950.1;  
 CC HSP; P30491; 1A1M.  
 CC MIM; 142830;  
 CC PROSITE; PS00290; IG\_MHC; 1.  
 CC PFAM; PF00047; ig; 1.  
 CC PFAM; PF00129; MHC.I; 1.  
 CC MHC.I; Transmembrane; Glycoprotein; Signal.  
 CC SIGNAL 1 24  
 CC CHAIN 25 362

FT DOMAIN 25 114 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,  
 FT DOMAIN 115 206 B\*49(B-21) B\*4901 ALPHA CHAIN.  
 FT DOMAIN 207 298 EXTRACELLULAR ALPHA-1.  
 FT DOMAIN 299 309 EXTRACELLULAR ALPHA-2.  
 FT TRANSMEM 310 333 EXTRACELLULAR ALPHA-3.  
 FT DOMAIN 334 362 CONNECTING PEPTIDE.  
 FT CARBOHYD 110 110 CYTOPLASMIC TAIL.  
 FT DISULFID 125 188 BY SIMILARITY.  
 FT DISULFID 227 283 BY SIMILARITY.  
 SQ SEQUENCE 362 AA; 40581 MW; E996F82F CRC32;

Query Match

45.0%; Score 45; DB 1; Length 362;  
 Best Local Similarity 62.5%; Pred. No. 4.3;

Matches 10; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 5 TRLNERENLRALRY 20  
 DB 93 TNTQTYRENLRALRY 108

RESULT 6  
 1B49 HUMAN  
 ID 1B49 HUMAN STANDARD; PRT; 362 AA.  
 AC P18464;  
 DT 01-NOV-1990 (Rel. 16, Created)  
 DT 01-NOV-1990 (Rel. 16, Last sequence update)  
 DT 01-FEB-1996 (Rel. 33, Last annotation update)  
 DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B\*51(B-5) B\*5101 ALPHA CHAIN  
 DE PRECURSOR.  
 GN HLA-B OR HLAB.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 90207291.  
 RA ENNIS P.D., ZEMMOUR J., SALTER R.D., PARHAM P.;  
 RT "Rapid cloning of HLA-A, B cDNA by using the polymerase chain  
 RT reaction: frequency and nature of errors produced in amplification.";  
 RN Proc. Natl. Acad. Sci. U.S.A. 87:2833-2837(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 89080265.  
 RA HAYASHI H., ENNIS P.D., ARIGA H., SALTER R.D., PARHAM P., KANO K.,  
 RA TAKIGUCHI M.;  
 RT "HLA-B\*51 and HLA-B\*52 differ by only two amino acids which are in the  
 RT helical region of the alpha 1 domain";  
 RL J. Immunol. 142:306-311(1989).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 89233295.  
 RA POHLA H., KUON W., TABACZEWSKI P., DOERNER C., WEISS E.H.;  
 RT "Allelic variation in HLA-B and HLA-C sequences and the evolution of  
 RT the HLA-B alleles";  
 RL Immunogenetics 29:297-307(1989).  
 CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO  
 CC THE IMMUNE SYSTEM.  
 CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
 CC MICROGLOBULIN).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; M32319; AAA36232.1; -  
 CC EMBL; M22792; AAA59620.1; ALT\_SEQ.  
 CC EMBL; M22786; AAA59620.1; JOINED.  
 CC EMBL; M22787; AAA59620.1; JOINED.  
 CC EMBL; M22788; AAA59620.1; JOINED.  
 CC EMBL; M22789; AAA59620.1; JOINED.  
 CC EMBL; M22790; AAA59620.1; JOINED.  
 CC EMBL; M22791; AAA59620.1; JOINED.  
 CC EMBL; L41087; AAA64513.1; -  
 CC EMBL; L41086; AAA64513.1; JOINED.  
 CC PIR; A30345; A30345.  
 CC PIR; A30548; A30548.  
 CC HSSP; P30491; IALM.  
 CC MIM; 142830; -  
 CC PROSITE; PS00290; IG\_MHC; 1.  
 CC PFAM; PF00047; Ig; 1.  
 CC MHC I; Transmembrane; Glycoprotein; Signal.

FT SIGNAL 1 24  
 FT CHAIN 25 362  
 FT  
 FT DOMAIN 25 114  
 FT DOMAIN 115 206  
 FT DOMAIN 207 298  
 FT DOMAIN 299 308  
 FT TRANSMEM 309 332  
 FT DOMAIN 333 362  
 FT CARBOHYD 110 110  
 FT DISULFID 125 188  
 FT DISULFID 227 283  
 SQ SEQUENCE 362 AA; 40566 MW; 4D846F30 CRC32;

Query Match 45.0%; Score 45; DB 1; Length 362;  
 Best Local Similarity 62.5%; Fred. No. 4.3;  
 Matches 10; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 5 TRLNERENLRALRY 20  
 DB 93 TNTQTYRENLRALRY 108

RESULT 7  
 1B52 HUMAN STANDARD; PRT; 362 AA.  
 ID 1B52 HUMAN STANDARD; PRT; 362 AA.  
 AC P30489;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 01-APR-1993 (Rel. 25, Last annotation update)  
 DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B\*51(B-5) B\*5104 ALPHA CHAIN  
 DE PRECURSOR.  
 GN HLA-B OR HLAB.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 92265955.  
 RA BELICH M.P., MADRIGAL J.A., HILDEBRAND W.H., ZEMMOUR J.,  
 RA WILLIAMS R.C., LOZ R., PETZL-ERLER M.L., PARHAM P.;  
 RT "Unusual HLA-B alleles in two tribes of Brazilian Indians.";  
 RL Nature 357:326-329(1992).  
 CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO  
 CC THE IMMUNE SYSTEM.  
 CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
 CC MICROGLOBULIN).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; Z15143; CAA78849.1; -  
 CC HSSP; P30491; IALM.  
 CC MIM; 142830; -  
 CC PROSITE; PS00290; IG\_MHC; 1.  
 CC PFAM; PF00047; Ig; 1.  
 CC PFAM; PF00129; MHC\_I; 1.  
 CC MHC I; Transmembrane; Glycoprotein; Signal.  
 FT SIGNAL 1 24  
 FT CHAIN 25 362  
 FT  
 FT DOMAIN 25 114  
 FT DOMAIN 115 206  
 FT DOMAIN 207 298  
 FT DOMAIN 299 308  
 FT TRANSMEM 309 332  
 FT DOMAIN 333 362

FT CARBOHYD 110 110 BY SIMILARITY.  
 FT DISULFID 125 188 BY SIMILARITY.  
 FT DISULFID 227 283 BY SIMILARITY.  
 SQ SEQUENCE 362 AA; 40560 MW; F22F08AB CRC32;

Query Match 45.08; Score 45; DB 1; Length 362;  
 Best Local Similarity 62.58; Pred. No. 4.3;  
 Matches 10; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 5 TRLNERRENLTALRY 20  
 I | | | | | | | |  
 Db 93 TINTOTYRENLTALRY 108

RESULT 8  
 1B53\_HUMAN  
 ID 1B53\_HUMAN STANDARD; PRT; 362 AA.  
 AC P30490;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 01-APR-1993 (Rel. 25, Last annotation update)  
 DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, BW-52(B-5) B\*5201 ALPHA CHAIN  
 DE PRECURSOR.  
 GN HLA-B OR HLAB.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 89080265.  
 RA HAYASHI H., ENNIS P.D., ARIGA H., SALTER R.D., PARHAM P., KANO K.,  
 RA TAKIGUCHI M.  
 RT "HLA-B51 and HLA-Bw52 differ by only two amino acids which are in the  
 RT helical region of the alpha 1 domain."  
 RL J. Immunol. 142:306-311(1989).  
 CC -1- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO  
 CC THE IMMUNE SYSTEM.  
 CC -1- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
 CC MICROGLOBULIN).

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----  
 CC EMBL; M22799; AAA59645.1; ALT SEQ.  
 CC EMBL; M22793; AAA59645.1; JOINED.  
 CC EMBL; M22794; AAA59645.1; JOINED.  
 CC EMBL; M22795; AAA59645.1; JOINED.  
 CC EMBL; M22796; AAA59645.1; JOINED.  
 CC EMBL; M22797; AAA59645.1; JOINED.  
 CC EMBL; M22798; AAA59645.1; JOINED.  
 CC PIR; B30345; B30345.  
 CC PIR; B30345; B30345.  
 CC PIR; B30548; B30548.  
 CC HSP; P30491; IAIM.  
 CC MIN; 142830.  
 CC PFAM; PF00047; Ig; 1.  
 CC PROSITE; PS00290; IG\_MHC; 1.  
 CC PFAM; PF00129; MHC.I; 1.  
 CC PFAM; PF00129; MHC.I; 1.  
 CC MHC I; Transmembrane; Glycoprotein; Signal.  
 CC SIGNAL 1 24  
 CC CHAIN 25 362 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,  
 CC DOMAIN 25 114 BW-52(B-5) B\*5201 ALPHA CHAIN.  
 CC DOMAIN 115 206 EXTRACELLULAR ALPHA-1.  
 CC DOMAIN 207 298 EXTRACELLULAR ALPHA-2.  
 CC DOMAIN 299 308 EXTRACELLULAR ALPHA-3.  
 CC DOMAIN 309 332 CONNECTING PEPTIDE.  
 CC DOMAIN 333 362 CYTOPLASMIC TAIL.

FT CARBOHYD 110 110 BY SIMILARITY.  
 FT DISULFID 125 188 BY SIMILARITY.  
 FT DISULFID 227 283 BY SIMILARITY.  
 SQ SEQUENCE 362 AA; 40521 MW; 3B436FEB CRC32;

Query Match 45.08; Score 45; DB 1; Length 362;  
 Best Local Similarity 62.58; Pred. No. 4.3;  
 Matches 10; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 5 TRLNERRENLTALRY 20  
 I | | | | | | | |  
 Db 93 TINTOTYRENLTALRY 108

RESULT 9  
 1B54\_HUMAN  
 ID 1B54\_HUMAN STANDARD; PRT; 362 AA.  
 AC P30491;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, BW-53 B\*5301 ALPHA CHAIN  
 DE PRECURSOR.  
 GN HLA-B OR HLAB.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 91033941.  
 RA HAYASHI H., Ooba T., NAKAYAMA S., SEKIMATA M., KANO K.,  
 RA TAKIGUCHI M.  
 RT "Allotopes of HLA-B\*53 differ by only two amino acids which are in the  
 RT substitution of the residues associated with HLA-Bw4/Bw6 public  
 RT epitopes."  
 RL Immunogenetics 32:195-199(1990).  
 CC -1- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO  
 CC THE IMMUNE SYSTEM.  
 CC -1- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
 CC MICROGLOBULIN).

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----  
 CC EMBL; M58636; AAA36228.1; -.  
 CC PIR; A45834; A45834.  
 CC PDB; 1AIM; 08-APR-98.  
 CC PDB; 1AIO; 08-APR-98.  
 CC MIN; 142830.  
 CC PROSITE; PS00290; IG\_MHC; 1.  
 CC PFAM; PF00047; Ig; 1.  
 CC PFAM; PF00129; MHC.I; 1.  
 CC MHC I; Transmembrane; Glycoprotein; Signal; 3D-structure.  
 CC SIGNAL 1 24  
 CC CHAIN 25 362 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,  
 CC DOMAIN 25 114 BW-53 B\*5301 ALPHA CHAIN.  
 CC DOMAIN 115 206 EXTRACELLULAR ALPHA-1.  
 CC DOMAIN 207 298 EXTRACELLULAR ALPHA-2.  
 CC DOMAIN 299 332 EXTRACELLULAR ALPHA-3.

FT DOMAIN 299 308 CONNECTING PEPTIDE.  
 FT TRANSMEM 309 332  
 FT DOMAIN 333 362 CYTOPLASMIC TAIL.  
 FT CARBOHYD 110 110 BY SIMILARITY.  
 FT DISULFID 125 188  
 FT DISULFID 227 283  
 SQ SEQUENCE 362 AA; 40495 MW; 2BDC746E CRC32;

Query Match 45.0%; Score 45; DB 1; Length 362;  
 Best Local Similarity 62.5%; Pred. No. 4.3;  
 Matches 10; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 5 TRLNERENRALTALRY 20  
 DB 93 TINTQTYRENRALTALRY 108

RESULT 10  
 BLH1\_YEAST STANDARD; PRT; 454 AA.  
 AC Q01532;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 15-DEC-1999 (Rel. 39, Last annotation update)  
 DE CYSTEINE PROTEINASE 1 (EC 3.4.22.-) (Y3) (BLEOMYCIN HYDROLASE) (BLM  
 DE HYDROLASE).  
 GN BLH1 OR YCPI OR LAP3 OR GAL6 OR YNL239W OR N1118.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Hemiascomycetes; Saccharomycetales;  
 OC Saccharomycetaceae; Saccharomycetes.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 93016103.  
 RA KAMBOURIS N.G., BURKE D.J., CREUTZ C.E.;  
 RT "Cloning and characterization of a cysteine proteinase from  
 RT Saccharomyces cerevisiae.";  
 RL J. Biol. Chem. 267:21570-21576(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 93144342.  
 RA MAGDOLEN U., MUELLER G., MAGDOLEN V., BANDLOW W.;  
 RT "A yeast gene (BLH1) encodes a polypeptide with high homology to  
 RT vertebrate bleomycin hydrolase, a family member of thiol  
 RT proteinases.";  
 RL Biochim. Biophys. Acta 1171:299-303(1993).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 93216638.  
 RA ENENKEL C., WOLF D.H.;  
 RT "BLH1 codes for a yeast thiol aminopeptidase, the equivalent of  
 RT mammalian bleomycin hydrolase.";  
 RL J. Biol. Chem. 268:7036-7043(1993).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 97051596.  
 RA PANDOLFO D., DE ANTONI A., LANFRANCHI G., VALLE G.;  
 RT "The DNA sequence of cosmid 14-5 from chromosome XIV reveals 21 open  
 RT reading frames including a novel gene encoding a globin-like  
 RT domain.";  
 RL Yeast 12:1071-1076(1996).  
 RN [5]  
 RP SEQUENCE OF 199-222 AND 338-358.  
 RX MEDLINE; 91353077.  
 RA CREUTZ C.E., SNYDER S.L., KAMBOURIS N.G.;  
 RT "Calcium-dependent secretory vesicle-binding and lipid-binding  
 RT proteins of Saccharomyces cerevisiae.";  
 RL Yeast 7:229-244(1991).  
 RN [6]  
 RP PARTIAL SEQUENCE FROM N.A., AND X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
 RX MEDLINE; 95365811.  
 RA JOSHUA-TOR L., XU H.E., JOHNSTON S.A., REES D.C.;  
 RT "Crystal structure of a conserved protease that binds DNA: the

RT bleomycin hydrolase, Gal6.";  
 RL Science 269:945-950(1995).  
 RN [7]  
 RP X-RAY CRYSTALLOGRAPHY (2.05 ANGSTROMS).  
 RX MEDLINE; 98206472.  
 RA ZHENG W., JOHNSTON S.A., JOSHUA-TOR L.;  
 RT "The unusual active site of Gal6/bleomycin hydrolase can act as a  
 RT carboxypeptidase, aminopeptidase, and peptide ligase.";  
 RL Cell 93:103-109(1998).  
 RN [8]  
 RP DNA-BINDING, AND CHARACTERIZATION.  
 RX MEDLINE; 94342285.  
 RA XU H.E., JOHNSTON S.A.;  
 RT "Yeast bleomycin hydrolase is a DNA-binding cysteine protease.  
 RT Identification, purification, biochemical characterization.";  
 RL J. Biol. Chem. 269:21177-21183(1994).  
 CC -!- FUNCTION: NON ESSENTIAL TO THE VIABILITY OF YEAST CELLS. CAN  
 CC PROTECT YEAST FROM THE TOXIC EFFECTS OF BLEOMYCIN. ACTS AS A  
 CC REPRESSOR IN THE GAL4 REGULATORY SYSTEM. BINDS SINGLE-STRANDED DNA  
 CC WITH HIGHER AFFINITY THAN DOUBLE-STRANDED DNA.  
 CC -!- COFACTOR: CALCIUM.  
 CC -!- SUBUNIT: HOMODIMER.  
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C1; ALSO KNOWN AS THE  
 CC PAPAIN FAMILY OF THIOL PROTEASES.  
 CC -!- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-30 IS THE INITIATOR,  
 CC ALTHOUGH FLANKING SEQUENCES POINT TO THE SECOND.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

CC EMBL; M97910; AAA35231.1; -  
 CC EMBL; X69124; CAA48878.1; -  
 CC EMBL; X68228; CAA48309.1; -  
 CC EMBL; Z71515; CAA96144.1; -  
 CC EMBL; Z69381; CAA93359.1; -  
 CC EMBL; U74299; AAB18260.1; -  
 CC PIR; S25606; S25606.  
 CC PIR; S28224; S28224.  
 CC PDB; 1GCB; 15-OCT-95.  
 CC PDB; 3GCB; 21-OCT-98.  
 CC PDB; 1A6R; 21-OCT-98.  
 CC SGD; L0000929; LAP3.  
 CC PROSITE; PS00139; THIOL\_PROTEASE\_CYS; 1.  
 CC PROSITE; PS00639; THIOL\_PROTEASE\_HIS; 1.  
 CC PROSITE; PS00640; THIOL\_PROTEASE\_ASN; FALSE.NEG.  
 KW Hydrolase; Thiol protease; DNA-binding; 3D-structure.  
 FT ACT\_SITE 73 73  
 FT ACT\_SITE 369 369  
 FT ACT\_SITE 392 392  
 FT REPEAT 28 48  
 FT REPEAT 193 213  
 FT CONFLICT 158 158 I -> M (IN REF. 2).  
 SQ SEQUENCE 454 AA; 52088 MW; B90FB48E CRC32;

Query Match 45.0%; Score 45; DB 1; Length 454;  
 Best Local Similarity 58.8%; Pred. No. 5.7;  
 Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 3 LATRLNERENRALTALR 19  
 DB 180 LTTKLREFAETLTALK 196

RESULT 11  
 1B01\_PANTR STANDARD; PRT; 359 AA.  
 ID 1B01\_PANTR



```

AC P13750;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 01-APR-1993 (Rel. 25, Last annotation update)
DE CHLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-1 ALPHA CHAIN PRECURSOR
DE (FRAGMENT).
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Pan.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 89030641.
RA KLEIN J.;
RA MAYER W.;
RA MAYER W.E., JONKER M., KLEIN D., IVANYI P., VAN SEVENTER G.,
RT "Nucleotide sequences of chimpanzee MHC class I alleles: evidence for
RT trans-species mode of evolution.";
RL EMBO J. 7:2765-2774(1988).
RN [2]
RP REVISIONS.
RA MAYER W.;
RL Submitted (FEB-1989) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC THE IMMUNE SYSTEM.
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC MICROGLOBULIN).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL: X13115; CAA31507.1;
DR PIR: S03537; S03537.
DR HSSP: P03989; IHSA.
DR PROSITE: PS00290; IG_MHC; 1.
DR PFAM: PF00047; IG; 1.
DR PFAM: PF00129; MHC_I; 1.
KW MHC I; Transmembrane; Glycoprotein; Signal.
FT NON_TER 1
FT SIGNAL <1 20
FT CHAIN 21 359
FT
FT CHLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
FT B-1 ALPHA CHAIN.
FT DOMAIN 21 110
FT DOMAIN 111 202
FT DOMAIN 203 294
FT DOMAIN 295 305
FT TRANSMEM 306 329
FT DOMAIN 330 359
FT DISULFID 121 184
FT DISULFID 223 279
FT CARBOHYD 106 106
FT SEQUENCE 359 AA; 40173 MW; 5395FFC9 CRC32;
SQ
Query Match 44.0%; Score 44; DB 1; Length 359;
Best Local Similarity 90.0%; Pred. No. 6.2;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Oy 11 RENLITARY 20
Db 95 RENLITARY 104
RESULT 12
ID 1B01_GORGO STANDARD; PRT; 362 AA.
AC P30379;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DE CHLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-1 ALPHA CHAIN PRECURSOR.
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 92078860.
RA LAWLOR D.A., WARREN E., TAYLOR P., PARHAM P.;
RT "Gorilla class I major histocompatibility complex alleles: comparison
RT to human and chimpanzee class I.";
J. Exp. Med. 174:1491-1509(1991).
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC THE IMMUNE SYSTEM.
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC MICROGLOBULIN).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL: X60255; CAA42807.1;
DR PIR: JH0539; JH0539.
DR HSSP: P03989; IHSA.
DR PROSITE: PS00290; IG_MHC; 1.
DR PFAM: PF00047; IG; 1.
DR PFAM: PF00129; MHC_I; 1.
KW MHC I; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 24
FT CHAIN 25 362
FT
FT CLASS I HISTOCOMPATIBILITY ANTIGEN.
FT GOGO-B0101 ALPHA CHAIN.
FT DOMAIN 25 114
FT DOMAIN 115 206
FT DOMAIN 207 298
FT DOMAIN 299 308
FT TRANSMEM 309 332
FT DOMAIN 333 362
FT DISULFID 125 188
FT DISULFID 227 283
FT CARBOHYD 110 110
FT SEQUENCE 362 AA; 40170 MW; 2E33E2B8 CRC32;
SQ
Query Match 44.0%; Score 44; DB 1; Length 362;
Best Local Similarity 90.0%; Pred. No. 6.3;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Oy 11 RENLITARY 20
Db 99 RENLITARY 108
RESULT 13
ID 1B02_GORGO STANDARD; PRT; 362 AA.
AC P30380;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DE CHLA CLASS I HISTOCOMPATIBILITY ANTIGEN, GOGO-B0102 ALPHA CHAIN PRECURSOR.
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 92078860.
RA LAWLOR D.A., WARREN E., TAYLOR P., PARHAM P.;
RT "Gorilla class I major histocompatibility complex alleles: comparison
RT to human and chimpanzee class I.";
J. Exp. Med. 174:1491-1509(1991).

```

```

DE OS CLASS I HISTOCOMPATIBILITY ANTIGEN, GOGO-B0101 ALPHA CHAIN PRECURSOR.
DE Gorilla gorilla gorilla (Lowland gorilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 92078860.
RA LAWLOR D.A., WARREN E., TAYLOR P., PARHAM P.;
RT "Gorilla class I major histocompatibility complex alleles: comparison
RT to human and chimpanzee class I.";
J. Exp. Med. 174:1491-1509(1991).
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC THE IMMUNE SYSTEM.
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC MICROGLOBULIN).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL: X60255; CAA42807.1;
DR PIR: JH0539; JH0539.
DR HSSP: P03989; IHSA.
DR PROSITE: PS00290; IG_MHC; 1.
DR PFAM: PF00047; IG; 1.
DR PFAM: PF00129; MHC_I; 1.
KW MHC I; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 24
FT CHAIN 25 362
FT
FT CLASS I HISTOCOMPATIBILITY ANTIGEN.
FT GOGO-B0101 ALPHA CHAIN.
FT DOMAIN 25 114
FT DOMAIN 115 206
FT DOMAIN 207 298
FT DOMAIN 299 308
FT TRANSMEM 309 332
FT DOMAIN 333 362
FT DISULFID 125 188
FT DISULFID 227 283
FT CARBOHYD 110 110
FT SEQUENCE 362 AA; 40170 MW; 2E33E2B8 CRC32;
SQ
Query Match 44.0%; Score 44; DB 1; Length 362;
Best Local Similarity 90.0%; Pred. No. 6.3;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Oy 11 RENLITARY 20
Db 99 RENLITARY 108
RESULT 13
ID 1B02_GORGO STANDARD; PRT; 362 AA.
AC P30380;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DE CHLA CLASS I HISTOCOMPATIBILITY ANTIGEN, GOGO-B0102 ALPHA CHAIN PRECURSOR.
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 92078860.
RA LAWLOR D.A., WARREN E., TAYLOR P., PARHAM P.;
RT "Gorilla class I major histocompatibility complex alleles: comparison
RT to human and chimpanzee class I.";
J. Exp. Med. 174:1491-1509(1991).

```



CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
DR EMBL; X03664; CAA27301.1; -;  
DR EMBL; X03667; CAA27301.1; JOINED.  
DR EMBL; L38504; ARA69724.1; -;  
DR PIR; B25092; HLHUBK.  
DR HSPF; P03989; LHSA.  
DR MIN; I42830; -;  
DR PROSITE; PS00290; IG\_MHC; 1.  
DR PFAM; PF00047; Ig; 1.  
DR PFAM; PF00129; MHC.I; 1.  
KW MHC I; Transmembrane; Glycoprotein; Signal.  
FT SIGNAL 1 24  
FT CHAIN 25 362 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,  
FT DOMAIN 25 114 B-27 B\*2702 ALPHA CHAIN.  
FT DOMAIN 115 206 EXTRACELLULAR ALPHA-1.  
FT DOMAIN 207 298 EXTRACELLULAR ALPHA-2.  
FT DOMAIN 299 308 EXTRACELLULAR ALPHA-3.  
FT TRANSMEM 309 332 CONNECTING PEPTIDE.  
FT DOMAIN 333 362 CYTOPLASMIC TAIL.  
FT CARBOHYD 110 110 BY SIMILARITY.  
FT DISULFID 125 188 BY SIMILARITY.  
FT DISULFID 227 283 BY SIMILARITY.  
SQ SEQUENCE 362 AA; 40397 MW; 9798F0BB CRC32;

Query Match 44.0%; Score 44; DB 1; Length 362;  
Best Local Similarity 90.0%; Pred. No. 6.3;  
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 11 RENTLALRY 20  
      |||||  
Db 99 RENTLALRY 108

Search completed: February 8, 2000, 01:26:00  
Job time: 1560 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2000, 19:16:14 : Search time 176.54 Seconds  
(without alignments)  
7.855 Million cell updates/sec

Title: US-08-653-294-30  
Perfect score: 100  
Sequence: 1 YRLATLNERENLTALRY 20

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 225878 seqs, 69334122 residues

Total number of hits satisfying chosen parameters: 225878

Minimum DB seq length: 0  
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database : SPTREMBL12:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mmc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	51	51.0	89	019674	019674 homo sapien
2	51	51.0	89	019565	019565 homo sapien
3	51	51.0	131	097998	097998 homo sapien
4	51	51.0	131	097999	097999 homo sapien
5	51	51.0	181	019779	019779 homo sapien
6	51	51.0	181	030197	030197 homo sapien
7	51	51.0	181	019669	019669 homo sapien
8	51	51.0	181	078028	078028 homo sapien
9	51	51.0	355	029853	029853 homo sapien
10	51	51.0	362	029637	029637 homo sapien
11	51	51.0	362	029935	029935 homo sapien
12	51	51.0	362	079524	079524 homo sapien
13	51	51.0	362	029850	029850 homo sapien
14	51	51.0	362	029661	029661 homo sapien
15	51	51.0	362	078180	078180 homo sapien
16	51	51.0	362	029933	029933 homo sapien
17	50	50.0	361	09XRY0	09XRY0 pongo pygma
18	48	48.0	137	095533	095533 pan troglod
19	46.5	46.5	346	095459	095459 rattus norv
20	46.5	46.5	346	078088	078088 rattus norv

21	46.5	46.5	348	7	046875	046875 rattus norv
22	46.5	46.5	362	7	030721	030721 macaca mula
23	46	46.0	89	4	095956	095956 homo sapien
24	46	46.0	89	7	077959	077959 homo sapien
25	46	46.0	181	7	077934	077934 homo sapien
26	46	46.0	330	7	019356	019356 macaca mula
27	46	46.0	331	7	002944	002944 macaca mula
28	46	46.0	331	7	002945	002945 macaca mula
29	46	46.0	362	7	029938	029938 homo sapien
30	46	46.0	368	7	061895	061895 mus musculus
31	45	45.0	89	7	019569	019569 homo sapien
32	45	45.0	172	7	019770	019770 homo sapien
33	45	45.0	172	7	019774	019774 homo sapien
34	45	45.0	172	7	019775	019775 homo sapien
35	45	45.0	172	7	019780	019780 homo sapien
36	45	45.0	172	7	095364	095364 homo sapien
37	45	45.0	172	7	019771	019771 homo sapien
38	45	45.0	172	7	019772	019772 homo sapien
39	45	45.0	172	7	019773	019773 homo sapien
40	45	45.0	175	7	029694	029694 homo sapien
41	45	45.0	181	7	046703	046703 homo sapien
42	45	45.0	181	7	062899	062899 homo sapien
43	45	45.0	181	7	062922	062922 homo sapien
44	45	45.0	181	7	019623	019623 homo sapien
45	45	45.0	181	7	019747	019747 homo sapien

ALIGNMENTS

RESULT 1  
019674 PRELIMINARY; PRT; 89 AA.  
AC 019674;  
DT 01-JAN-1998 (TREMREL. 05, Created)  
DT 01-JAN-1998 (TREMREL. 05, Last sequence update)  
DT 01-NOV-1998 (TREMREL. 08, Last annotation update)  
DE HLA-B\*13 (FRAGMENT).  
GN HLA-B\*13.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA MAERTENS R., DE CANCK I.;  
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; Y12378; CAA73021.1; -.  
DR PFAM; PF00129; MHC\_I; 1.  
KW MHC.  
FT NON\_TER 1 1  
FT NON\_TER 89 89  
SQ SEQUENCE 89 AA; 10581 MW; 9AC7631C CRC32;

Query Match 51.0%; Score 51; DB 7; Length 89;  
Best Local Similarity 68.8%; Pred. No. 0.59;  
Matches 11; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 TRLNERENLTALRY 20  
| | | | | | | | | |  
DB 68 TNTQTYRENLTALRY 83

RESULT 2  
019565 PRELIMINARY; PRT; 89 AA.  
ID 019565;  
AC 019565;  
DT 01-JAN-1998 (TREMREL. 05, Created)  
DT 01-MAY-1999 (TREMREL. 10, Last sequence update)  
DT 01-MAY-1999 (TREMREL. 10, Last annotation update)  
DE MHC CLASS I ANTIGEN (FRAGMENT).  
GN HLA-B.  
OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA CAO K., BURDETT L., ZHANG G., FERNANDEZ-VINA M.;  
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF017316; AAB70282.2; -  
 KW MHC  
 FT NON\_TER 1 1  
 FT NON\_TER 89 89  
 SQ SEQUENCE 89 AA; 10581 MW; 9AC7631C CRC32;

Query Match 51.0%; Score 51; DB 7; Length 89;  
 Best Local Similarity 68.8%; Pred. No. 0.59;  
 Matches 11; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 5 TRLNERRENLTALRY 20  
 DB 68 TNTQTYRENLTALRY 83

RESULT 3  
 O97998  
 ID O97998 PRELIMINARY; PRT; 131 AA.  
 AC O97998;  
 DT 01-MAY-1999 (TREMBLrel. 10, Created)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)  
 DE MHC CLASS I ANTIGEN (FRAGMENT).  
 GN HLA-B.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA HOLDSWORTH R.;  
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF035648; AAD02035.1; -  
 KW MHC  
 FT NON\_TER 1 1  
 FT NON\_TER 131 131  
 SQ SEQUENCE 131 AA; 15134 MW; 931C8D27 CRC32;

Query Match 51.0%; Score 51; DB 7; Length 131;  
 Best Local Similarity 68.8%; Pred. No. 0.87;  
 Matches 11; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 5 TRLNERRENLTALRY 20  
 DB 34 TNTQTYRENLTALRY 49

RESULT 4  
 O97999  
 ID O97999 PRELIMINARY; PRT; 131 AA.  
 AC O97999;  
 DT 01-MAY-1999 (TREMBLrel. 10, Created)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)  
 DE MHC CLASS I ANTIGEN (FRAGMENT).  
 GN HLA-B.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA HOLDSWORTH R.;  
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF035649; AAD02036.1; -  
 KW MHC  
 FT NON\_TER 1 1

FT NON\_TER 131 131  
 SQ SEQUENCE 131 AA; 15057 MW; 32C865A3 CRC32;

Query Match 51.0%; Score 51; DB 7; Length 131;  
 Best Local Similarity 68.8%; Pred. No. 0.87;  
 Matches 11; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 5 TRLNERRENLTALRY 20  
 DB 34 TNTQTYRENLTALRY 49

RESULT 5  
 O19779  
 ID O19779 PRELIMINARY; PRT; 181 AA.  
 AC O19779;  
 DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
 DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)  
 DE MHC CLASS I ANTIGEN (FRAGMENT).  
 GN HLA-B.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STRINER N.K., NG J., BUSH J., HARTZMAN R.J., JOHNSON-DOW L.,  
 RA HURLEY C.K.;  
 RL Hum. Immunol. 56:0-0(1997).  
 DR EMBL; U90241; AAB82305.1; -  
 DR EMBL; U90240; AAB82305.1; JOINED.  
 DR HSSP; P10318; 1ROG.  
 DR PFAM; PF00129; MHC\_I; 1.  
 KW MHC.  
 FT NON\_TER 1 1  
 FT NON\_TER 181 181  
 SQ SEQUENCE 181 AA; 21095 MW; 97EC2597 CRC32;

Query Match 51.0%; Score 51; DB 7; Length 181;  
 Best Local Similarity 68.8%; Pred. No. 1.2;  
 Matches 11; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 5 TRLNERRENLTALRY 20  
 DB 68 TNTQTYRENLTALRY 83

RESULT 6  
 O30197  
 ID O30197 PRELIMINARY; PRT; 181 AA.  
 AC O30197;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)  
 DE MHC CLASS I ANTIGEN (FRAGMENT).  
 GN HLA-B.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA HURLEY C.K., HOVER R.J.;  
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U63560; AAB05925.1; -  
 DR EMBL; U63559; AAB05925.1; JOINED.  
 DR HSSP; P10318; 1ROG.  
 DR PFAM; PF00129; MHC\_I; 1.  
 KW MHC.  
 FT NON\_TER 1 1  
 FT NON\_TER 181 181  
 SQ SEQUENCE 181 AA; 20997 MW; 70FABE37 CRC32;

Query Match 51.0%; Score 51; DB 7; Length 181;  
 Best Local Similarity 68.8%; Pred. No. 1.2;  
 Matches 11; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 TRLNERENLRTALRY 20  
 | |||||  
 Db 68 TNTQTYENLRTALRY 83

RESULT 7  
 Q19669 PRELIMINARY; PRT; 181 AA.  
 AC Q19669;  
 DT 01-JAN-1998 (TEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TEMBLrel. 05, Last sequence update)  
 DT 01-NOV-1999 (TEMBLrel. 12, Last annotation update)  
 DE MHC CLASS I ANTIGEN SHCHA (FRAGMENT).  
 GN HLA-B.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA HURLEY C.K., STEINER N.K.;  
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U58470; AAB02607.1; -.  
 DR EMBL; U58469; AAB02607.1; JOINED.  
 DR HSSP; P10318; 1ROG.  
 DR PFAM; PF00129; MHC\_I; 1.  
 KW MHC.  
 FT NON\_TER 1  
 FT NON\_TER 181  
 SQ SEQUENCE 181 AA; 21028 MW; EC872642 CRC32;

Query Match 51.0%; Score 51; DB 7; Length 181;  
 Best Local Similarity 68.8%; Pred. No. 1.2;  
 Matches 11; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 TRLNERENLRTALRY 20  
 | |||||  
 Db 68 TNTQTYENLRTALRY 83

RESULT 8  
 Q78028 PRELIMINARY; PRT; 181 AA.  
 AC Q78028;  
 DT 01-NOV-1998 (TEMBLrel. 08, Created)  
 DT 01-NOV-1998 (TEMBLrel. 08, Last sequence update)  
 DT 01-NOV-1999 (TEMBLrel. 12, Last annotation update)  
 DE HLA-B\*15MD (FRAGMENT).  
 GN HLA-B\*15MD.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-BLOOD;  
 RA GAO X., MATHESON B.;  
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U58316; AAB87723.1; -.  
 DR EMBL; U58315; AAB87723.1; JOINED.  
 DR HSSP; P10318; 1ROG.  
 DR PFAM; PF00129; MHC\_I; 1.  
 KW MHC.  
 FT NON\_TER 1  
 FT NON\_TER 181  
 SQ SEQUENCE 181 AA; 21029 MW; C35A18BE CRC32;

Query Match 51.0%; Score 51; DB 7; Length 181;  
 Best Local Similarity 68.8%; Pred. No. 1.2;  
 Matches 11; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 TRLNERENLRTALRY 20  
 | |||||  
 Db 68 TNTQTYENLRTALRY 83

RESULT 9  
 Q29853 PRELIMINARY; PRT; 355 AA.  
 AC Q29853;  
 DT 01-NOV-1996 (TEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)  
 DT 01-NOV-1999 (TEMBLrel. 12, Last annotation update)  
 DE HLA-B ALPHA-CHAIN (FRAGMENT).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE; 95176328.  
 RA GAUCHAT-FEISS D., BREUR-VRIESENDORP B.S., RUFER N., JEANNET M.,  
 RA ROOSNEK E., TIERCY J.M.;  
 RT "Sequencing of a novel functional HLA-B\*44 subtype differing in two  
 RT residues in the alpha 2 domain";  
 RL Tissue Antigens 44:261-264(1994).  
 DR EMBL; X75953; CAA53566.1; -.  
 DR HSSP; P30491; 1ALM.  
 DR PROSITE; PS00290; IG\_MHC; 1.  
 DR PFAM; PF00047; Ig; 1.  
 DR PFAM; PF00129; MHC\_I; 1.  
 KW MHC.  
 FT NON\_TER 355  
 FT NON\_TER 355  
 SQ SEQUENCE 355 AA; 39845 MW; EF81934E CRC32;

Query Match 51.0%; Score 51; DB 7; Length 355;  
 Best Local Similarity 68.8%; Pred. No. 2.4;  
 Matches 11; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 TRLNERENLRTALRY 20  
 | |||||  
 Db 93 TNTQTYENLRTALRY 108

RESULT 10  
 Q29637 PRELIMINARY; PRT; 362 AA.  
 AC Q29637;  
 DT 01-NOV-1996 (TEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)  
 DT 01-NOV-1999 (TEMBLrel. 12, Last annotation update)  
 DE MHC CLASS I ANTIGEN.  
 GN HLA-B.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA DOMENA J.D., PARHAM P.;  
 RL Submitted (DEC-1993) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U04244; AAB87397.1; -.  
 DR HSSP; P30460; 1AGE.  
 DR PROSITE; PS00290; IG\_MHC; 1.  
 DR PFAM; PF00047; Ig; 1.  
 DR PFAM; PF00129; MHC\_I; 1.  
 KW MHC.  
 SQ SEQUENCE 362 AA; 40359 MW; EEAB95D7 CRC32;

Query Match 51.0%; Score 51; DB 7; Length 362;





Query Match 51.0%; Score 51; DB 7; Length 362;  
 Best Local Similarity 68.8%; Pred. No. 2.4;  
 Matches 11; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 TRLNERRENLTALRY 20  
 | | | | | | | | | |  
 Db 93 TNTQTYRENLTALRY 108

## RESULT 15

O78180 PRELIMINARY; PRT; 362 AA.  
 AC O78180;  
 DT 01-NOV-1998 (TREMELrel. 08, Created)  
 DT 01-NOV-1998 (TREMELrel. 08, Last sequence update)  
 DT 01-NOV-1999 (TREMELrel. 12, Last annotation update)  
 DE HLA CLASS I ANTIGEN HLA-B.  
 GN HLA-B.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA ELLEXSON M.E., HILDEBRAND W.H.;  
 RT "Molecular characterization of HLA-B\*1303new."  
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U75533; AAC31793.1; -  
 DR HSSP; F30491; IAHM.  
 DR PROSITE; PS00290; IG\_MHC; 1.  
 DR PFAM; PF00047; Ig; 1.  
 DR PFAM; PF00129; MHC\_I; 1.  
 KW MHC.  
 SQ SEQUENCE 362 AA; 40476 MW; 8BA369FA CRC32;

Query Match 51.0%; Score 51; DB 7; Length 362;  
 Best Local Similarity 68.8%; Pred. No. 2.4;  
 Matches 11; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 TRLNERRENLTALRY 20  
 | | | | | | | | | |  
 Db 93 TNTQTYRENLTALRY 108

Search completed: February 8, 2000, 19:16:14  
 Job time: 21501 sec

**THIS PAGE BLANK (USPTO)**



```

* 12168 13182: contig of 1015 bp in length
*      gap of unknown length
* 13183 14040: contig of 858 bp in length
*      gap of unknown length
* 14041 15296: contig of 1256 bp in length
*      gap of unknown length
* 15297 16215: contig of 919 bp in length
*      gap of unknown length
* 16216 17301: contig of 1086 bp in length
*      gap of unknown length
* 17302 18060: contig of 759 bp in length
*      gap of unknown length
* 18061 18903: contig of 843 bp in length
*      gap of unknown length
* 18904 19771: contig of 868 bp in length
*      gap of unknown length
* 19772 20568: contig of 797 bp in length
*      gap of unknown length
* 20569 21622: contig of 1054 bp in length
*      gap of unknown length
* 21623 22449: contig of 827 bp in length
*      gap of unknown length
* 22450 23873: contig of 1424 bp in length
*      gap of unknown length
* 23874 25949: contig of 2076 bp in length.
FEATURES      Location/Qualifiers
Source        1..25949
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /chromosome="5"
              /clone="CITB-H1.2168D21"
BASE COUNT   7182 a 5843 c 5600 g 7316 t      8 others
ORIGIN

```

```

alignment_scores:
  Quality: 52.00      Length: 15
  Ratio: 4.000       Gaps: 0
Percent Similarity: 86.667 Percent Identity: 73.333

```

```

alignment_block:
US-08-653-294-30 x AC010413

```

```

Align seg 1/1 to: AC010413 from: 1 to: 25949

```

```

2 ArgLeuAlaThrArgLeuAsnGluArgGluAsnLeuArgThr 16
||||| ||||||||| ||||||| ||||||| |||||||
8826 AAATCTCAACACGCTGCAACATCAGAGGAACAACCTCCGGACA 8870

```

```

seq_name: gb_htg5:AC015865

```

```

seq_documentation_block:
LOCUS      AC015865 123347 bp      DNA      HTG      17-NOV-1999
DEFINITION Homo sapiens clone RP11-55A1, LOW-PASS SEQUENCE SAMPLING.
ACCESSION  AC015865
VERSION     AC015865.1 GI:6446803
KEYWORDS   HTG: HTGS_PHASE0; NULL.
SOURCE      human.

```

```

ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 123347)
AUTHORS     Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE       Homo sapiens, clone RP11-55A1
JOURNAL     Unpublished
REFERENCE   2 (bases 1 to 123347)
AUTHORS     Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
            Baldwin,J., Barna,N., Beckerly,R., Boguslavsky,L., Boukhgalter,B.,
            Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
            Cooke,P., Dearellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
            Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D., Horton,L.,
            Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,

```

```

Lehoczyk, J., Lieu, C., Locke, K., Macdonald, P., Marquis, N.,
McEwan, P., McGurk, A., McKernan, K., McLaughlin, J., Meldrim, J.,
Morrow, J., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P.,
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Testaye, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X.,
Wyman, D., Ye, W.J., Zimmer, A. and Zody, M.
Direct Submission
Submitted (17-NOV-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Project Information
Center project name: L626
Center clone name: 55_A_1
-----

```

```

* NOTE: This record contains 140 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

```

```

1
* 868: contig of 868 bp in length
*      gap of unknown length
* 869 1706: contig of 838 bp in length
*      gap of unknown length
* 1707 2529: contig of 823 bp in length
*      gap of unknown length
* 2530 3392: contig of 863 bp in length
*      gap of unknown length
* 3393 4216: contig of 824 bp in length
*      gap of unknown length
* 4217 5052: contig of 836 bp in length
*      gap of unknown length
* 5053 5879: contig of 827 bp in length
*      gap of unknown length
* 5880 6713: contig of 834 bp in length
*      gap of unknown length
* 6714 7550: contig of 837 bp in length
*      gap of unknown length
* 7551 8392: contig of 842 bp in length
*      gap of unknown length
* 8393 9234: contig of 842 bp in length
*      gap of unknown length
* 9235 10079: contig of 845 bp in length
*      gap of unknown length
* 10080 10934: contig of 855 bp in length
*      gap of unknown length
* 10935 11767: contig of 833 bp in length
*      gap of unknown length
* 11768 12598: contig of 831 bp in length
*      gap of unknown length
* 12599 13443: contig of 845 bp in length
*      gap of unknown length
* 13444 14281: contig of 838 bp in length
*      gap of unknown length
* 14282 15101: contig of 820 bp in length
*      gap of unknown length
* 15102 15950: contig of 849 bp in length
*      gap of unknown length
* 15951 16782: contig of 832 bp in length
*      gap of unknown length
* 16783 17627: contig of 845 bp in length
*      gap of unknown length
* 17628 18473: contig of 846 bp in length
*      gap of unknown length
* 18474 19302: contig of 829 bp in length

```

```

TITLE
JOURNAL
COMMENT

```

```
* *
* * 19303 20156: contig of 854 bp in length
* * gap of unknown length
* * 20157 20994: contig of 838 bp in length
* * gap of unknown length
* * 20995 21914: contig of 820 bp in length
* * gap of unknown length
* * 21815 22644: contig of 830 bp in length
* * gap of unknown length
* * 22645 23466: contig of 822 bp in length
* * gap of unknown length
* * 23467 24313: contig of 847 bp in length
* * gap of unknown length
* * 24314 25156: contig of 843 bp in length
* * gap of unknown length
* * 25157 25981: contig of 825 bp in length
* * gap of unknown length
* * 25982 26819: contig of 838 bp in length
* * gap of unknown length
* * 26820 27671: contig of 852 bp in length
* * gap of unknown length
* * 27672 28514: contig of 843 bp in length
* * gap of unknown length
* * 28515 29448: contig of 934 bp in length
* * gap of unknown length
* * 29449 30281: contig of 833 bp in length
* * gap of unknown length
* * 30282 31107: contig of 826 bp in length
* * gap of unknown length
* * 31108 31948: contig of 841 bp in length
* * gap of unknown length
* * 31949 32786: contig of 838 bp in length
* * gap of unknown length
* * 32787 33641: contig of 855 bp in length
* * gap of unknown length
* * 33642 34485: contig of 844 bp in length
* * gap of unknown length
* * 34486 35305: contig of 820 bp in length
* * gap of unknown length
* * 35306 36153: contig of 848 bp in length
* * gap of unknown length
* * 36154 36993: contig of 840 bp in length
* * gap of unknown length
* * 36994 37824: contig of 831 bp in length
* * gap of unknown length
* * 37825 38653: contig of 829 bp in length
* * gap of unknown length
* * 38654 39490: contig of 837 bp in length
* * gap of unknown length
* * 39491 40345: contig of 855 bp in length
* * gap of unknown length
* * 40346 41178: contig of 833 bp in length
* * gap of unknown length
* * 41179 42015: contig of 837 bp in length
* * gap of unknown length
* * 42016 42860: contig of 845 bp in length
* * gap of unknown length
* * 42861 43696: contig of 836 bp in length
* * gap of unknown length
* * 43697 44524: contig of 828 bp in length
* * gap of unknown length
* * 44525 45347: contig of 823 bp in length
* * gap of unknown length
* * 45348 46177: contig of 830 bp in length
* * gap of unknown length
* * 46178 47013: contig of 836 bp in length
* * gap of unknown length
* * 47014 47836: contig of 823 bp in length
* * gap of unknown length
* * 47837 48670: contig of 834 bp in length
* * gap of unknown length
* * 48671 49502: contig of 832 bp in length
* * gap of unknown length
```

```
* *
* * 49503 50363: contig of 861 bp in length
* * gap of unknown length
* * 50364 51194: contig of 831 bp in length
* * gap of unknown length
* * 51195 52027: contig of 833 bp in length
* * gap of unknown length
* * 52028 52875: contig of 848 bp in length
* * gap of unknown length
* * 52876 53712: contig of 837 bp in length
* * gap of unknown length
* * 53713 54539: contig of 827 bp in length
* * gap of unknown length
* * 54540 55376: contig of 837 bp in length
* * gap of unknown length
* * 55377 56207: contig of 831 bp in length
* * gap of unknown length
* * 56208 57066: contig of 859 bp in length
* * gap of unknown length
* * 57067 57881: contig of 815 bp in length
* * gap of unknown length
* * 57882 58715: contig of 834 bp in length
* * gap of unknown length
* * 58716 59549: contig of 834 bp in length
* * gap of unknown length
* * 59550 60371: contig of 822 bp in length
* * gap of unknown length
* * 60372 61213: contig of 842 bp in length
* * gap of unknown length
* * 61214 62061: contig of 848 bp in length
* * gap of unknown length
* * 62062 62993: contig of 932 bp in length
* * gap of unknown length
* * 62994 63913: contig of 920 bp in length
```

## alignment\_scores:

```
Quality: 52.00 Length: 15
Ratio: 4.000 Gaps: 0
Percent Similarity: 86.667 Percent Identity: 73.333
```

## alignment\_block:

```
US-08-653-294-30 x AC015865 ..
```

```
Align seg 1/1 to: AC015865 from: 1 to: 123347
```

```
2 ArgLeuAlaThrArgLeuAsnGluArgGluAsnLeuArgThr 16
:::||||| ||||| ||||| ||||| ::||| ||||| |||||
47640 AAATCCGACACGCTGTGATATCATAGAGGACACCTCCGGACA 47684
```

```
seq_name: gb_htg4:AC011290
```

## seq\_documentation\_block:

```
LOCUS AC011290 148409 bp DNA HTG 15-OCT-1999
DEFINITION Homo sapiens clone NH0064102, *** SEQUENCING IN PROGRESS ***, 3
```

```
unordered pieces.
```

```
ACCESSION AC011290
```

```
VERSION AC011290.2 GI:6042094
```

```
KEYWORDS HTG: HTGS_PHASE1.
```

```
SOURCE human.
```

```
ORGANISM Homo sapiens
```

```
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
```

```
Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

```
REFERENCE 1 (bases 1 to 148409)
```

```
AUTHORS Waterston,R.H.
```

```
TITLE The sequence of Homo sapiens clone
```

```
JOURNAL Unpublished
```

```
REFERENCE 2 (bases 1 to 148409)
```

```
AUTHORS Waterston,R.H.
```

```
TITLE Direct Submission
```

```
JOURNAL Submitted (05-OCT-1999) Genome Sequencing Center, Washington
```

```
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
```

```
MO 63108, USA
```

```
COMMENT On Oct 15, 1999 this sequence version replaced gi:6012159.
```

```
* NOTE: This is a 'working draft' sequence. It currently
```

\* consists of 3 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

\* 1 1991: contig of 1991 bp in length  
 \* 1992 2009: gap of unknown length  
 \* 2010 64050: contig of 62041 bp in length  
 \* 64051 84088: gap of unknown length  
 \* 64069 148409: contig of 84341 bp in length.

## FEATURES

source  
 1. 148409  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="NH0064102"

BASE COUNT 44805 a 31502 c 29724 g 42342 t 36 others  
 ORIGIN

alignment\_scores:  
 Quality: 52.00 Length: 15  
 Ratio: 4.000 Gaps: 0  
 Percent Similarity: 86.667 Percent Identity: 73.333

## alignment\_block:

US-08-653-294-30 x AC011290 ..

Align seg 1/1 to: AC011290 from: 1 to: 148409

2 ArgLeuAlaThrArgLeuAsnGluArgGluAsnLeuAlaGThr 16

33983 AAACCTCCGACACGCTGTAATATCATGAGGACACCTCCGGACA 34027

seq\_name: gb\_htg4:AC010352

## seq\_documentation\_block:

LOCUS AC010352 153027 bp DNA HTG 31-OCT-1999  
 DEFINITION Homo sapiens chromosome 5 clone C17B-HL\_2022F20, \*\*\* SEQUENCING IN  
 PROGRESS \*\*\*, 77 unordered pieces.

ACCESSION AC010352  
 VERSION AC010352.2 GI:6165069  
 KEYWORDS HTG; HTGS\_PHASE1.  
 SOURCE human.

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 153027)

AUTHORS DOE Joint Genome Institute.

TITLE Sequencing of Human Chromosome 5

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 153027)

DOE Joint Genome Institute.

AUTHORS Direct Submission

TITLE Submitted (15-SEP-1999), Production Sequencing Facility, DOE Joint

JOURNAL Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

COMMENT On Oct 31, 1999 this sequence version replaced gi:5882582.

\* NOTE: This is a 'working draft' sequence. It currently

\* consists of 77 contigs. The true order of the pieces

\* is not known and their order in this sequence record is

\* arbitrary. Gaps between the contigs are represented as

\* runs of N, but the exact sizes of the gaps are unknown.

\* This record will be updated with the finished sequence

\* as soon as it is available and the accession number will

\* be preserved.

\* 1 675: contig of 675 bp in length

\* 676 gap of unknown length

\* 1399: contig of 724 bp in length

\* gap of unknown length

\* 1400 2124: contig of 725 bp in length

\* gap of unknown length

\* 2125 2891: contig of 767 bp in length

\* 2892 gap of unknown length

\* 3576: contig of 685 bp in length

\* gap of unknown length

\* 3577 4442: contig of 866 bp in length

\* gap of unknown length

\* 4443 5226: contig of 784 bp in length

\* gap of unknown length

\* 5227 5881: contig of 655 bp in length

\* gap of unknown length

\* 5882 6591: contig of 710 bp in length

\* gap of unknown length

\* 6592 7225: contig of 634 bp in length

\* gap of unknown length

\* 7226 7780: contig of 555 bp in length

\* gap of unknown length

\* 7781 8459: contig of 679 bp in length

\* gap of unknown length

\* 8460 9189: contig of 730 bp in length

\* gap of unknown length

\* 9190 9894: contig of 705 bp in length

\* gap of unknown length

\* 9895 10528: contig of 634 bp in length

\* gap of unknown length

\* 10529 11169: contig of 641 bp in length

\* gap of unknown length

\* 11170 11816: contig of 647 bp in length

\* gap of unknown length

\* 11817 12641: contig of 825 bp in length

\* gap of unknown length

\* 12642 13127: contig of 486 bp in length

\* gap of unknown length

\* 13128 13960: contig of 833 bp in length

\* gap of unknown length

\* 13961 14620: contig of 660 bp in length

\* gap of unknown length

\* 14621 15321: contig of 701 bp in length

\* gap of unknown length

\* 15322 16382: contig of 1061 bp in length

\* gap of unknown length

\* 16383 16563: contig of 181 bp in length

\* gap of unknown length

\* 16564 17592: contig of 1029 bp in length

\* gap of unknown length

\* 17593 18581: contig of 989 bp in length

\* gap of unknown length

\* 18582 19292: contig of 711 bp in length

\* gap of unknown length

\* 19293 20208: contig of 916 bp in length

\* gap of unknown length

\* 20209 21122: contig of 914 bp in length

\* gap of unknown length

\* 21123 21499: contig of 377 bp in length

\* gap of unknown length

\* 21500 22611: contig of 1112 bp in length

\* gap of unknown length

\* 22612 23418: contig of 807 bp in length

\* gap of unknown length

\* 23419 24584: contig of 1166 bp in length

\* gap of unknown length

\* 24585 25024: contig of 440 bp in length

\* gap of unknown length

\* 25025 25800: contig of 776 bp in length

\* gap of unknown length

\* 25801 26106: contig of 306 bp in length

\* gap of unknown length

\* 26107 26994: contig of 888 bp in length

\* gap of unknown length

\* 26995 27627: contig of 633 bp in length

\* gap of unknown length

\* 27628 28281: contig of 654 bp in length

\* gap of unknown length

\* 28282 29585: contig of 1304 bp in length

\* gap of unknown length

```

FEATURES                                  Location/Qualifiers
    source                                1..153027
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /chromosome="5"
        /clone="CIRB-HI_2022F20"
BASE COUNT      43128 a 33071 c 33103 g 43621 t   104 others
ORIGIN

alignment_scores:
    Quality:             52.00           Length:          15
    Ratio:               4.000           Gaps:            0
Percent Similarity:     86.667         Percent Identity: 73.333

alignment_block:
US-08-653-294-30 x AC010352/rev ..

Align seg 1/1 to reverse of: AC010352 from: 1 to: 153027

      2 ArgLeuAlaThrArgLeuAsnGluArgArgGluAsnLeuArgThr 16
      :::||| ||||||| ||||||| ||||||| ::|||
135173 AAATCTCAACAGCGTGTGCACATCAGAAGACACCTCCGGACA 135129

seq_name: gb_pri:HSB44HLA2

seq_documentation_block:
LOCUS       HSB44HLA2              163 bp      DNA                PRI
DEFINITION  H.sapiens HLA-B gene, exon 2.
ACCESSION   X78849
VERSION     X78849.1 GI:602874
KEYWORDS    heavy chain; MHC class I; MHC class I HLA-B*44.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominoidea; Homo.

REFERENCE   1 (bases 1 to 163)
AUTHORS    Yao,Z., Volgger,A., Scholz,S., Bonisch,J. and Albert,E.D.
TITLE      Nucleotide sequence of a novel HLA-B*44 subtype B*4405
JOURNAL    Immunogenetics 40 (4), 310 (1994)
MEDLINE    94364642
REFERENCE   2 (bases 1 to 163)
AUTHORS    Albert,E.
DIRECT SUBMISSION
TITLE      Submitted (11-APR-1994) E. Albert, Labor fuer Immungenetik,
JOURNAL    Kinderpoliklinik der LMU Muenchen, Pettenhoferstr. 8a, 80336
Muenchen, FRG

FEATURES                                  Location/Qualifiers
    source                                1..163
        /organism="Homo sapiens"
        /isolate="WJG"
        /db_xref="taxon:9606"
        /chromosome="6"
        /germline
        /note="MHC"
exon                                             1..163
        /note="MHC Class I heavy chain HLA-B*44 variant"
        /number=2
BASE COUNT      45 a 50 c 54 g 14 t
ORIGIN

alignment_scores:
    Quality:             51.00           Length:          16
    Ratio:               3.923           Gaps:            0
Percent Similarity:     81.250         Percent Identity: 68.750

alignment_block:
US-08-653-294-30 x HSB44HLA2 ..

Align seg 1/1 to: HSB44HLA2 from: 1 to: 163

      5 ThrArgLeuAsnGluArgArgGluAsnLeuArgThrAlaLeuArgTyr 20

```







/label-HLA-B*44_CDS			
/product="human leucocyte antigen B"			
/protein_id="CAB56342.2"			
/db_xref="GI:6015538"			
/translation="SHSMRYFYAMSRGRGEPRFTIVGVYDDTLFVREDSDATSPRK"			
EPRAPWIEQSGPEYWDRETOISKNTQTYENLTALRYNQSEA"			
BASE COUNT	60 a	89 c	85 g
ORIGIN	36 t		
alignment_scores:			
Quality:	51.00	Length:	16
Ratio:	3.923	Gaps:	0
Percent Similarity:	81.250	Percent Identity:	68.750
alignment_block:			
US-08-653-294-30 x HSA249724 ..			
Align seg 1/1 to: HSA249724 from: 1 to: 270			
5 ThrArgLeuAsnGluArgArgGluAsnLeuArgThrAlaLeuArgTyr 20			
: : :			
204 ACCACACACAGACTTACCGAGAGAACCTGCGACCGCGCTCCGCTAC 251			
seq_name: gb_pr2:HSSHCHA01			
seq_documentation_block:			
LOCUS	HSSHCHA01	270 bp	DNA PRI 18-JUN-1996
DEFINITION	Human MHC class I antigen SHCHA (HLA-B*4403 variant) gene, exon 2.		
ACCESSION	U58469		
VERSION	U58469.1 GI:1378136		
KEYWORDS	1 of 2		
SEGMENT	human.		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;		
REFERENCE	Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 270)		
TITLE	Hurley,C.K. and Steiner,N.K.		
JOURNAL	Novel HLA-B alleles		
REFERENCE	Unpublished		
AUTHORS	2 (bases 1 to 270)		
TITLE	Hurley,C.K. and Steiner,N.K.		
JOURNAL	Direct Submission		
FEATURES	Submitted (16-MAY-1996) C.K. Hurley, Microbiology & Immunology,		
source	Georgetown University, 3970 Reservoir Rd.NW, Washington, DC 20007,		
Location/Qualifiers	USA		
1..270			
/organism="Homo sapiens"			
/db_xref="taxon:9606"			
/cell_type="PBL"			
/haplotype="B*5101/B*4403 variant"			
/chromosome="6"			
/note="SHCHA-B*4403 variant, Direct sequence"			
1..270			
/gene="HLA-B"			
/note="HLA-B*4403 variant"			
/number=2			
60 a	89 c	85 g	36 t
BASE COUNT			
ORIGIN			
alignment_scores:			
Quality:	51.00	Length:	16
Ratio:	3.923	Gaps:	0
Percent Similarity:	81.250	Percent Identity:	68.750
alignment_block:			
US-08-653-294-30 x HSSHCHA01 ..			
Align seg 1/1 to: HSSHCHA01 from: 1 to: 270			

```

5 ThrArgLeuAsnGluArgArgGluAsnLeuAlaLeuArgTyr 20
||||: ::: |||||||||||||||||||||||
204 ACCAACACACAGACTTACCGAGAGAACCTGCGCACCGCGCTCGCTAC 251

seq_name: gb_pr3: HSHLAB15M1

seq_documentation_block:
LOCUS HSHLAB15M1 270 bp DNA PRI 01-DEC-1997
DEFINITION Human HLA-B*15MD gene, exon 2.
ACCESSION U58315
VERSION U58315.1 GI:2654406
KEYWORDS
SEGMENT
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Gao,X. and Matheson,B.
TITLE A novel b15 variant found in oceanic populations
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 270)
AUTHORS Gao,X.
TITLE Direct Submission
JOURNAL Submitted (15-MAY-1996) Xiaojiang Gao, The Australian National
University, Human Genetics Group, Acton, Canberra, Act, Australia.
2601

FEATURES
source Location/Qualifiers
1..270
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/map="6p21"
/cell_line="MD674"
/tissue_type="lymphocyte"
1..270
/gene="HLA-B*15MD"
/number=2
58 a 91 c 85 g 36 t
BASE COUNT
ORIGIN

alignment_scores:
Quality: 51.00 Length: 16
Ratio: 3.923 Gaps: 0
Percent Similarity: 81.250 Percent Identity: 68.750

alignment_block:
US-08-653-294-30 x HSHLAB15M1 ..
Align seg 1/1 to: HSHLAB15M1 from: 1 to: 270

5 ThrArgLeuAsnGluArgArgGluAsnLeuAlaLeuArgTyr 20
||||: ::: |||||||||||||||||||||||
204 ACCAACACACAGACTTACCGAGAGAACCTGCGCACCGCGCTCGCTAC 251

seq_name: gb_pr3: HSHLABSH1

seq_documentation_block:
LOCUS HSHLABSH1 270 bp DNA PRI 30-OCT-1997
DEFINITION Human MHC class I antigen HLA-B GN00155-B*38022 gene, exon 2.
ACCESSION U90240
VERSION U90240.1 GI:1905853
KEYWORDS
SEGMENT
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Steiner,N.K., Ng,J., Bush,J., Hartzman,R.J., Johnson-Dow,L. and
Hurley,C.K.

```

•

```

TITLE      HLA-B Alleles Associated with the B15 Serologically Defined
JOURNAL    Antigen
REFERENCE   Hum. Immunol. 56 (1997) In press
AUTHORS    Steiner, N.K. and Hurley, C.K.
TITLE      Direct Submission
JOURNAL    Submitted (20-FEB-1997) Microbiology & Immunology, Georgetown
           University, 3970 Reservoir Rd. NW, Washington, DC 20007, USA
FEATURES   Location/Qualifiers
            source      1..270
                        /organism="Homo sapiens"
                        /isolate="GN00155"
                        /db_xref="taxon:9606"
            exon        1..270
                        /gene="HLA-B"
                        /number=2
BASE COUNT   60 a 88 c 85 g 37 t
ORIGIN
alignment_scores:
  Quality: 51.00      Length: 16
  Ratio: 3.923       Gaps: 0
Percent Similarity: 81.250 Percent Identity: 68.750

alignment_block:
US-08-653-294-30 x HSHLABSH1 ..
Align seg 1/1 to: HSHLABSH1 from: 1 to: 270

5 ThrArgLeuAsnGluArgGluAsnLeuArgThrAlaLeuArgTyr 20
||||: ::|||
204 ACCAACACACAGACTTACCGAGAGACCTGCGCAGCGCGCTCGCTAC 251

seq_name: gb_pr4:AF017316

seq_documentation_block:
LOCUS      AF017316      270 bp      DNA      PRI      26-MAR-1999
DEFINITION Homo sapiens MHC class I antigen (HLA-B) gene, partial cds.
ACCESSION  AF017316
VERSION    AF017316.1 GI:2394335
KEYWORDS   human.
SOURCE     Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
           Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 270)
AUTHORS   Cao, K., Burdett, L., Zhang, G. and Fernandez-Vina, M.
TITLE     Direct Submission
JOURNAL   Submitted (07-AUG-1997) Nat. Histocompatibility Lab, Am Red Cross,
           box 173, 22 S. Greene St., Baltimore, MD 21201, USA
FEATURES   Location/Qualifiers
            source      1..270
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /clone="27B"
                        /chromosome="6"
                        <1..>270
                        /gene="HLA-B"
                        /note="novel allele"
                        <1..>270
                        /gene="HLA-B"
                        /codon_start=3
                        /product="MHC class I antigen"
                        /protein_id="AAB70282.2"
                        /db_xref="GI:4521330"
                        /translation="SHSMRVFYTAMSRPGRPFITGVGVDDTFQVRFDSATSPRM
                        APRAPWIEQGEPEVMDRETQISKTNTQTYRENLTALRYNQSEA"
BASE COUNT   59 a 92 c 83 g 36 t
ORIGIN
alignment_scores:
  Quality: 51.00      Length: 16
  Ratio: 3.923       Gaps: 0
Percent Similarity: 81.250 Percent Identity: 68.750

alignment_block:
US-08-653-294-30 x AF017316 ..
Align seg 1/1 to: AF017316 from: 1 to: 270

5 ThrArgLeuAsnGluArgGluAsnLeuArgThrAlaLeuArgTyr 20
||||: ::|||
204 ACCAACACACAGACTTACCGAGAGACCTGCGCAGCGCGCTCGCTAC 251

seq_name: gb_pr4:AF017316

seq_documentation_block:
LOCUS      AF017316      270 bp      DNA      PRI      23-SEP-1999
DEFINITION Homo sapiens HLA-B MHC class I antigen (HLA-B) gene, HLA-B-*38021
ACCESSION  AF181857
VERSION    AF181857.1 GI:5919134
KEYWORDS   human.
SOURCE     Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
           Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 270)
AUTHORS   Pimantothai, N. and Hurley, C.K.
TITLE     Novel HLA-B allele
JOURNAL   Unpublished
REFERENCE  2 (bases 1 to 270)
AUTHORS   Pimantothai, N. and Hurley, C.K.
TITLE     Direct Submission
JOURNAL   Submitted (30-AUG-1999) Microbiology & Immunology, Georgetown
           University, 3970 Reservoir Rd. N.W., Washington, D.C. 20007, USA
FEATURES   Location/Qualifiers
            source      1..270
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /clone="49N-TA"
                        <1..>270
                        /gene="HLA-B"
                        /number=2
            exon        58 a 87 c 88 g 37 t
BASE COUNT   58 a 87 c 88 g 37 t
ORIGIN
alignment_scores:
  Quality: 51.00      Length: 16
  Ratio: 3.923       Gaps: 0
Percent Similarity: 81.250 Percent Identity: 68.750

alignment_block:
US-08-653-294-30 x HB38021G1 ..
Align seg 1/1 to: HB38021G1 from: 1 to: 270

5 ThrArgLeuAsnGluArgGluAsnLeuArgThrAlaLeuArgTyr 20
||||: ::|||
204 ACCAACACACAGACTTACCGAGAGACCTGCGCAGCGCGCTCGCTAC 251

```

**THIS PAGE BLANK (USPTO)**

Sequence	Strd	Orig	Zscore	EScore	Len	Documentation
N_Genesed_36:T42608	+	46.00	137.82	6.37	285	! Synthetic gene, LqHV encoding
N_Genesed_36:T029167	+	45.00	135.21	4.25	270	! HLA-B*52 exon 2 alpha-1 domain
N_Genesed_36:T029168	+	45.00	135.21	4.25	270	! HLA-B*52 exon 2 alpha-1 domain
N_Genesed_36:T001834	+	45.00	122.32	32.65	1086	! Sequence encoding HLA-B*51 anti
N_Genesed_36:T001834	+	45.00	122.32	32.65	1086	! Sequence encoding HLA-B*52 anti
N_Genesed_36:T001834	+	45.00	122.32	32.65	1086	! Sequence encoding HLA-B*52 anti
N_Genesed_36:T005693	+	45.00	122.29	32.75	1089	! HLA-B*51 gene for production of
N_Genesed_36:T005701	+	45.00	122.29	32.75	1089	! HLA-B*51 gene for production of
N_Genesed_36:T012114	+	45.00	122.29	32.75	1089	! HLA-B*53 exon. HLA-B*53 gene,
N_Genesed_36:T020545	+	45.00	122.29	32.75	1089	! HLA-B*53 exon. HLA-B*53 gene,
N_Genesed_36:T020545	+	45.00	111.64	128.49	3442	! Encodes KEX2 endopeptidase with
N_Genesed_36:T098696	+	45.00	110.60	158.44	4106	! Yeast KEX2 gene. Yeast with re
N_Genesed_36:T025236	+	44.00	104.08	338.44	5558	! Streptococcus pneumoniae genom
N_Genesed_36:T054249	-	44.00	100.12	562.75	8528	! Construct EC2L (Contains catR
N_Genesed_36:T041225	+	43.50	111.15	136.71	2190	! Clone PC14 encoding major surf
N_Genesed_36:T07447	+	43.00	126.26	119.68	362	! B. subtilis aprE target DNA. De
N_Genesed_36:T080745	+	43.00	115.01	83.32	1220	! Coding region of the aprA gene
N_Genesed_36:T080745	+	43.00	115.01	83.32	1220	! Coding region of the aprA gene
N_Genesed_36:T00587	-	43.00	115.01	83.32	1220	! Thermotable Bacillus subtilin.
N_Genesed_36:T00890	-	43.00	115.01	83.32	1220	! Thermotable Bacillus subtilin.
N_Genesed_36:T003536	-	43.00	113.11	106.41	1499	! Subtilisin gene. Mutant prokar
N_Genesed_36:T071241	+	43.00	113.10	106.49	1500	! Subtilisin gene from Bacillus
N_Genesed_36:T00058	-	43.00	113.10	106.49	1500	! Bacillus subtilis subtilisin g
N_Genesed_36:T090042	-	43.00	113.10	106.49	1500	! Bacillus subtilis subtilisin g
N_Genesed_36:T090042	-	43.00	113.10	106.49	1500	! Bacillus subtilis subtilisin g
N_Genesed_36:T060475	+	43.00	112.95	108.52	1524	! Sequence of the apr [Bsu] gene
N_Genesed_36:T060475	+	43.00	112.95	108.52	1524	! Sequence of the apr [Bsu] gene
N_Genesed_36:T060475	+	43.00	107.45	219.71	2760	! Bacillus subtilis metalloprote
N_Genesed_36:T060475	+	43.00	107.45	219.71	2760	! Bacillus subtilis metalloprote
N_Genesed_36:T03487	+	43.00	104.77	309.90	3687	! DNA for Bacillus stearothermo
N_Genesed_36:T008695	+	43.00	104.41	324.43	3832	! sbSA gene encoding S-layer pro
N_Genesed_36:T028343	+	42.00	111.00	139.35	1344	! PBULI core sequence from Lact
N_Genesed_36:T028343	+	42.00	111.00	139.35	1344	! PBULI core sequence from Lact
N_Genesed_36:T049262	+	42.00	102.85	364.95	3023	! FcHV-B gag/pol gene. Multi-val
N_Genesed_36:T049262	+	42.00	102.85	364.95	3023	! FcHV-B gag/pol gene. Multi-val
N_Genesed_36:T082261	+	42.00	102.85	396.71	3243	! Enkephalinase gene (rat). Deco
N_Genesed_36:T082261	+	42.00	102.85	396.71	3243	! Enkephalinase gene (rat). Deco
N_Genesed_36:T062179	+	42.00	102.85	396.71	3243	! Rat enkephalinase cDNA. Metho
N_Genesed_36:T062179	+	42.00	102.85	396.71	3243	! Rat enkephalinase cDNA. Metho
N_Genesed_36:T029114	+	42.00	101.69	460.09	3674	! FeLV-A gag/pol fragment. Vacci
N_Genesed_36:T029114	+	42.00	101.69	460.09	3674	! FeLV-A gag/pol fragment. Vacci
N_Genesed_36:T013265	+	42.00	93.99	1.2e+03	8440	! FeLV F6A provirus clone 61E ga
N_Genesed_36:T020550	+	42.00	93.99	1.2e+03	8440	! FeLV F6A provirus clone 61E ga
N_Genesed_36:T020550	+	42.00	93.99	1.2e+03	8440	! FeLV F6A provirus clone 61E ga
N_Genesed_36:T060544	+	41.50	122.04	33.84	345	! Polynucleotide sequence from t
N_Genesed_36:T060544	+	41.50	122.04	33.84	345	! Polynucleotide sequence from t
N_Genesed_36:T060544	+	41.50	122.04	33.84	345	! Polynucleotide sequence from t
N_Genesed_36:T039144	+	41.00	136.23	5.48	6	









CC native promoter of the gene was removed and replaced with the  
CC *Aspergillus* glucoamylase promoter gene. This modification allows  
CC increased expression of the *catr* gene without the need to supply  
CC hydrogen peroxide to induce expression. Cells into which this  
CC construct is inserted preferably have the glucose oxidase gene  
CC (*goxA*) deleted. This deletion minimises the generation of  
CC gluconate waste material and the use of waste treatment processes.  
SQ Sequence 8528 BP; 1951 A; 2321 C; 2209 G; 2047 T;

```
alignment_scores:
  Quality: 44.00
  Ratio: 3.143
  Percent Similarity: 82.353
  Percent Identity: 52.941
  Gaps: 0
  Length: 17
```

alignment\_block:  
US-08-653-294-30 x Q46249/rev

Align seg 1/1 to reverse of: Q46249 from: 1 to: 8528

1776 TATCGCTTAGCATCTAGCTTGAATTCCTGCAGAGACCATCTCGCCACTTC 1727

17 a 17

name: N\_Geneseq\_36:04

seq\_name: N\_Geneseq\_36:Q41225

AC Q46249;

02-SEP-1993 (first entry)  
Clone PC14 encoding major surface gp of rat P. carini.  
Major surface glycoprotein; gp116; rat; Pneumocystis carinii;  
vaccine; HIV; human immunodeficiency virus; diagnostic; PCR;  
Rat Pneumocystis carinii.

FT misc\_feature 1. .8

FT 5'utr 9..1957

FT	cds	1958. .4370
----	-----	-------------

FT 3'utr 4371. 4432

```
FT
/*tag= e
```

```
FT      poly_site
4441.  .6093
/**** f
```

FT signals

```

F.F
F.F
F.F
F.F
/*tag= g
/*tag= g

```

$$\frac{h}{\hbar} = 2\pi$$

PN W09318166-A.

04-MAR-1993; 002020.  
04-MAR-1992: US-845989.

PA (GEMV) GENENCOR INT INC

DR WPI; 93-303480/38.  
PT *Aspergillus niger* catp

PT been inserted

CC The Aspergillus niger ca

```

SQ sequence 302 BP, 59 A, 63 C, 62 G, 92 I,
alignment_scores:
  Quality: 43.00      Length: 17
  Ratio: 2.867       Gaps: 0
  Percent Similarity: 88.235  Percent Identity: 52.941
alignment_block:
  US-08-653-294-30 x T07447/rev ..
  Align seg 1/1 to reverse of: T07447 from: 1 to: 362
      3 LeuAlaThrArgLeuAsnGluArgGluAsnLeuArgThrAlaLeuAr 19
      381 GTAGCCTTGAGAGTGAAGCGCGCTTTAAATTGAGAAATGCCATAAG 332
      19 g 19
      331 G 61
  seq_name: N_Geneseq_36:N80745
  seq_documentation_block:
    ID N80745 standard; DNA; 1220 BP.
    AC N80745;
    DE aprA gene encoding subtilin.
    DT 18-FEB-1991 (first entry)
    DR DE
alignment_scores:
  Quality: 43.00      Length: 17
  Ratio: 2.867       Gaps: 0
  Percent Similarity: 88.235  Percent Identity: 52.941
alignment_block:
  US-08-653-294-30 x N80745/rev ..
  Align seg 1/1 to reverse of: N80745 from: 1 to: 1220
      3 LeuAlaThrArgLeuAsnGluArgGluAsnLeuArgThrAlaLeuAr 19
      381 GTAGCCTTGAGAGTGAAGCGCGCTTTAAATTGAGAAATGCCATAAG 332
      19 g 19
      331 G 331
  seq_name: N_Geneseq_36:Q06587
  seq_documentation_block:
    ID Q06587 standard; DNA; 1220 BP.
    AC Q06587;
    DT 18-FEB-1991 (first entry)
    DR DE
    DE aprA gene encoding subtilin.

```

KW Detergents; depilatory tanning; serology; ds.  
OS Bacillus subtilis.  
FH Key Location/Qualifiers  
FT conflict 568..573  
FT /tag= a  
FT /label= gtagc  
FT /note= "Stahl et al. J.Bacteriol.,158,411-418 (1984)"  
FT cds 1..1143  
FT /tag= b  
FT mat\_peptide 319..1143  
FT /tag= c  
PN EP-398539-A.  
PD 22-NOV-1990.  
PF 01-MAY-1990: 304715.  
PR 17-MAY-1989: US-353124.  
PA (AMGE-) AMGEN INC.  
PI Zukowski MM, Narih IO, Levitt M;  
DR WPI; 90-350298/47.  
DR P-PSDB; R07970.  
PT Bacillus subtilin analogues - with improved pH thermal and oxidn.  
PT stability useful in cleaning compns.  
PS Claim 20; Table 1; 39pp; English.  
CC Modified analogues of subtilin are useful in cleaning fabrics, and  
CC have an improved resistance to oxidation, heat and pH extremes.  
CC Analogues have one or more negative AAs present in the calcium  
CC binding site (CBS) and may also be used in manufacture of protein  
CC hydrolsates and detection of incomplete Abs in serology.  
SQ Sequence 1220 BP; 355 A; 281 C; 283 G; 301 T;

alignment\_scores:  
Quality: 43.00 Length: 17  
Ratio: 2.867 Gaps: 0  
Percent Similarity: 88.235 Percent Identity: 52.941

alignment\_block:  
US-08-653-294-30 x Q06587/rev ...  
Align seg 1/1 to reverse of: Q06587 from: 1 to: 1220  
3 LeuAlaThrArgLeuAsnGluArgArgGluAsnLeuArgThrAlaLeuAr 19  
:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |  
381 GTAGCCTTGAGAGTGAAGCGCGCTTTAATTGAGAAATGCCATAAG 332  
19 q 19  
331 G 331

**THIS PAGE BLANK (USPTO)**

Sequence	Strd Orig	ZScore	EScore	Len	Document
gb_est33:AI124815	+	51.00	147.60	18.65	AI124815
gb_gss4:AO720034	+	50.00	142.11	37.75	HS_5541_A2_G07_SP6E RH
gb_gss15:AO625497	+	48.00	136.87	73.89	AIQ220034
gb_est1:DS35316	+	47.00	136.69	75.59	AIQ255497
gb_est1:DS35321	+	47.00	136.69	75.59	CELK019GGF
gb_est1:DS35329	+	47.00	136.69	75.59	CELK019GGF
gb_est1:DS35379	+	47.00	136.69	75.59	CELK022GGF
gb_est17:CG9731	+	47.00	136.68	75.59	CELK022GGF
gb_est17:CG9775	+	47.00	136.68	75.59	CELK022GGF
gb_est17:CG9958	+	47.00	136.69	75.59	CELK022GGF
gb_est17:DS37592	+	47.00	135.69	85.90	CELK004CAF
gb_est31:AI676428	+	47.00	133.52	113.45	ETMEST04183
gb_est28:AI488819	+	47.00	131.01	156.72	EST247158
gb_est36:AV186046	+	46.00	131.62	112.02	AV186046
gb_est33:AI777709	+	46.00	132.77	124.93	EST235804
gb_est33:AI774937	+	46.00	130.52	166.87	EST235804
gb_est34:AI808160	+	46.00	130.10	176.13	W53309.3
gb_gss4:AO713859	+	46.00	129.00	202.72	AI713859
gb_est32:AI738072	+	46.00	127.18	256.14	BSBML33Z484SK
gb_est8:CO3945	+	45.00	134.82	96.11	CO3945
gb_est11:RA263158	+	45.00	132.89	123.06	PMY0534
gb_est40:AA263668	+	45.00	130.91	158.58	PMY0534
gb_est6:BG82221	+	45.00	130.16	174.65	HUMHBC4626
gb_est6:BG3040	+	45.00	129.32	194.56	TgESTy59b10
gb_gss6:AO640208	+	45.00	127.93	232.35	AO640208
gb_gss6:AO849043	+	45.00	127.61	242.35	AO849043
gb_gss3:CB22809	+	45.00	127.45	247.30	CB22809
gb_gss34:AO572791	+	45.00	126.81	268.56	F2F8FR
gb_est36:AO001511	+	45.00	126.17	291.44	AO001511
gb_est20:AA881004	+	45.00	126.03	296.57	AA881004
gb_est38:AW092336	+	45.00	125.29	326.41	AW092336
gb_gss15:AO655562	+	45.00	125.18	331.00	AO655562
gb_est9:AA106132	+	45.00	125.08	334.95	AA106132
gb_gss40:AA254983	+	45.00	124.56	358.14	mgx50007C3dr
gb_gss9:AO083696	+	45.00	124.36	367.50	CG05818
gb_est31:AI696864	+	45.00	123.47	412.24	AI696864
gb_gss3:BI2131	-	45.00	119.85	655.48	BI2131
gb_est21:AA980664	-	44.00	136.72	75.32	AA980664
gb_est20:AA897217	+	44.00	132.76	125.09	AA897217
gb_est10:AA522865	+	44.00	130.87	159.43	AA522865
gb_est10:AA518891	+	44.00	130.83	160.21	AA518891
gb_est33:AI776145	+	44.00	130.65	164.12	AI776145
gb_gss1:FR0004928	+	44.00	129.13	199.35	FR0004928

```

seq_documentation_block:
LOCUS   A0720034          532 bp      DNA          GSS          14-JUL-1999
DEFINITION   HS.5541_A2.G07.SP6E.RPCI-11 Human Male BAC Library Homo sapiens
              genomic clone Plate-1117 Col-14 Row-W, genomic survey sequence.
ACCESSION   A0720034
VERSION     A0720034.1 GI:5479703
KEYWORDS    GSS.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 532)
AUTHORS    Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
TITLE      Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
JOURNAL    Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
MEDLINE    99380589
COMMENT    On Feb 19, 1999 this sequence version replaced gi:4142976.
            Contact: Mahairas GG, Wallace JC, Hood L
            High throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering.bac.htm)
            or from Resear h Genetics (info@resgen.com). BAC end web Server:
            http://www.htsc.washington.edu
            Plate: 1117 row: M column: 14
            Seq primer: SP6
            Class: BAC ends
            High quality sequence stop: 532.
FEATURES             Location/Qualifiers
     source           1..532
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="Plate-1117 Col-14 Row-W"
                     /clone_lib="RPCI-11 Human Male BAC Library"
                     /sex="male"
                     /note="vector: pBACE3.6; Genomic sequence of BAC ends"
BASE COUNT   145 a 119 c 104 g 158 t 6 others
ORIGIN
alignment_scores:
  Quality: 50.00      Length: 16
  Ratio: 3.571       Gaps: 0
  Percent Similarity: 87.500      Percent Identity: 62.500
alignment_block:
US-08-653-294-30 x A0720034
..
Align seg 1/1 to: A0720034 from: 1 to: 532
1 TyArgLeuAlaThrArgLeuAsnGluArgArgGluAsnLeuArgThr 16
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
433 TATAAGTGGCTTTTACTCAATGCTAGCGAGAGAGGTTGCATACA 480

seq_name: gb_gss15:A0625497

seq_documentation_block:
LOCUS   A0625497          485 bp      DNA          GSS          16-JUN-1999
DEFINITION   CITBI-EI-2653J22.TF CITBI-EI Homo sapiens genomic clone 2653J22,
              genomic survey sequence.
ACCESSION   A0625497
VERSION     A0625497.1 GI:5087889
KEYWORDS    GSS.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 485)
AUTHORS    Zhao,S., Adams,M.D., Nierman,W., Malek,J., Shizuya,H., Simon,M. and
            Venter,J.C.
TITLE      Use of BAC End Sequences from CalTech Libraries for Sequence-Ready
            Map Building
JOURNAL    Unpublished (1997)
COMMENT    Other_GSSs: CITBI-EI-2653J22.TR
            Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbe@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13-21
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source           1..485
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="2653J22"
                     /clone_lib="CITBI-EI"
                     /sex="male"
                     /cell_type="sperm"
                     /note="vector: pBelOBAC11; Site_1: EcoRI; Site_2: EcoRI;
                     Caltech Human BAC Library D"
BASE COUNT   181 a 76 c 80 g 148 t
ORIGIN
alignment_scores:
  Quality: 48.00      Length: 19
  Ratio: 3.200       Gaps: 0
  Percent Similarity: 78.947      Percent Identity: 52.632
alignment_block:
US-08-653-294-30 x A0625497
..
Align seg 1/1 to: A0625497 from: 1 to: 485
2 ArgLeuAlaThrArgLeuAsnGluArgArgGluAsnLeuArgThrAlaLe 18
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
108 AGATGCTCAACATTACTAAATGTTAGGAAATGAAATCAAAATCGCACT 157

seq_name: gb_est1:D35316

seq_documentation_block:
LOCUS   D35316            360 bp      mRNA          EST          08-AUG-1994
DEFINITION   CELK019GZF Yuji Kohara unpublished cDNA Caenorhabditis elegans cdNA
              clone yk19g12 5', mRNA sequence.
ACCESSION   D35316
VERSION     D35316.1 GI:526832
KEYWORDS    EST.
SOURCE      Caenorhabditis elegans.
            Caenorhabditis elegans.
            Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
            Rhabditina; Rhabditidae; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE   1 (bases 1 to 360)
AUTHORS    Kohara,Y., Mitsuaki,H., Nishigaki,A., Motohashi,T., Sugimoto,A. and
            Tabara,H.
TITLE      Toward an expression map of the C.elegans genome
JOURNAL    Unpublished (1994)
COMMENT    Contact: Yuji Kohara

```

```

BASE COUNT      116 a      85 c      79 g      80 t
ORIGIN

alignment_scores:
  Quality:      47.00      Length:      17
  Ratio:        3.615      Gaps:      0
  Percent Similarity: 76.471      Percent Identity: 64.706

alignment_block:
  US-08-653-294-30 x D35321      ..

Align seg 1/1 to: D35321 from: 1 to: 360

2 ArgLeuAlaThrArgLeuAsnGluArgGluAsnLeuArgThrAlaLe 18
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
79 AGATCCGGGTTAAAGGTTGACGGAAAGAGATCCAATCTGAGGATTATCT 128

```

129 G 129

seq\_name: gb\_est1:D35579

seq\_documentation\_block:

LOCUS D35579 360 bp mRNA EST 08-AUG-1994

DEFINITION CELK02259F Yuji Kohara unpublished cDNA Caenorhabditis elegans cDNA clone YK2299 5', mRNA sequence.

ACCESSION D35579

VERSION D35579.1 GI:527030

KEYWORDS EST.

SOURCE Caenorhabditis elegans.

ORGANISM Caenorhabditis elegans

Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida; Rhabditina; Rhabditoidae; Rhabditidae; Peloderinae; Caenorhabditis.

REFERENCE 1 (bases 1 to 360)

AUTHORS Kohara, Y., Mitsuki, H., Nishizaki, A., Motohashi, T., Sugimoto, A. and

```

TITLE      Toward an expression map of the C.elegans genome
JOURNAL    Unpublished (1994)
COMMENT    Contact: Yuji Kohara
           Gene Library Lab
           National Institute of Genetics
           Yata 1111, Mishima, Shizuoka 411, Japan
           Tel: 81-559-81-6854
           Fax: 81-559-81-6855
           Email: ykohara@lab.nig.ac.jp.

FEATURES   Location/Qualifiers
            1..360
             /organism="Caenorhabditis elegans"
             /strain="CB1489 him-8(e1489)"
             /db_xref="taxon:6239"
             /clone="YK22g9"
             /clone_lib="Yuji Kohara unpublished cDNA"
             /sex="hermaphrodite, male"
             /tissue_type="whole animal"
             /dev_stage="varied"
BASE COUNT 106 a      84 c      87 g      77 t      6 others
ORIGIN

alignment_scores:
Quality:          47.00      Length:          17
Ratio:            3.615     Gaps:              0
Percent Similarity: 76.471   Percent Identity: 64.706

alignment_block:
US-08-653-294-30 x D35579 ..

Align seg 1/1 to: D35579 from: 1 to: 360

```

ACCESSION C69775  
VERSION C69775.1 GI:2440300  
KEYWORDS EST.

Fax: 01-559-01-0055





US-08-653-294-30 x AI676428

Align seg 1/1 to: AI676428 from: 1 to: 499

1 TyrArgLeuAlaThrArgLeuAsnGluArgGluAsnLeuArgThr 16  
|||||  
1 TATCGACTTCCACGATGCGATCAGGAGGAGGATGAGGACT 48

seq\_name: gb\_est28:AI488819

seq\_documentation\_block: 647 bp mRNA EST 29-JUN-1999  
LOCUS AI488819  
DEFINITION EST247158 tomato ovary, TAMU Lycopersicon esculentum cDNA clone  
CLES1818, mRNA sequence.  
ACCESSION AI488819  
VERSION AI488819.1 GI:4384190  
KEYWORDS EST.  
SOURCE tomato.  
ORGANISM Lycopersicon esculentum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core  
eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum;  
Potatoe; Lycopersicon.  
REFERENCE 1 (bases 1 to 647)  
AUTHORS Alcala,J., Vrebalov,J., White,R., Matern,A.L., Vision,T.,  
Holt,I.E., Liang,F., Upton,J., Ronning,C.M., Craven,M.B.,  
Fuji,C.Y., Bowman,C.L., Niernan,W., Fraser,C.M., Venter,J.C.,  
Martin,G.B., Tanksley,S.D. and Giovannoni,J.  
Generation of ESTs from tomato carpel tissue  
Unpublished (1999)  
TITLE On May 18, 1998 this sequence version replaced gi:3138293.  
JOURNAL  
COMMENT  
Contact: David Frisch  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Tel: 864 656 4366  
Fax: 864 656 4293  
Email: dfrisch@CLEMSON.EDU.

FEATURES  
source  
Location/Qualifiers  
1..647

/organism="Lycopersicon esculentum"  
/cultivar="TA496"  
/db\_xref="taxon:4081"  
/clone="cLED1818"  
/clone\_lib="tomato ovary, TAMU"  
/tissue\_type="carpel"  
/dev\_stage="5.days pre-anthesis to 5 days post-anthesis"  
/lab\_host="XL1-Blue MRF"  
/note="Vector: pBlueScript SK(-); Site\_1: EcoRI; Site\_2:  
XhoI; cLED - Tomato Carpel EST Library. OligodT-primed and  
directionally cloned cDNA in vector Lambda ZAP II with 5'  
and 3' ends located at the EcoRI and XhoI sites,  
respectively."

BASE COUNT 198 a 116 c 148 g 185 t  
ORIGIN

alignment\_scores:  
Quality: 47.00 Length: 20  
Ratio: 3.615 Gaps: 0  
Percent Similarity: 65.000 Percent Identity: 55.000

alignment\_block:

US-08-653-294-30 x AI488819/rev ..

Align seg 1/1 to reverse of: AI488819 from: 1 to: 647

1 TyrArgLeuAlaThrArgLeuAsnGluArgGluAsnLeuArgThrAl 17  
|||||  
384 TATGTCGCAATACGATTCGACGATTCGGACAAATTCAGGACAG 335

17 aLeuArgTyr 20  
|||||

334 GTTGAGGTAT 325

seq\_name: gb\_est36:AV186046

seq\_documentation\_block: 360 bp mRNA EST 22-JUL-1999  
LOCUS AV186046  
DEFINITION AV186046 Yuji Kohara unpublished cDNA:Strain N2 hermaphrodite  
embryo Caenorhabditis elegans cDNA clone yk4946 5', mRNA sequence.  
ACCESSION AV186046  
VERSION AV186046.1 GI:5568029  
KEYWORDS EST.  
SOURCE Caenorhabditis elegans.  
ORGANISM Caenorhabditis elegans.  
Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditidae;  
Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
REFERENCE 1 (bases 1 to 360)  
AUTHORS Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Mitsuki,H.,  
Nishigaki,A., Motohashi,T., Zeng,Q., Watanabe,H., Sugimoto,A.,  
Sano,M., Miyata,A., Mitani,Y., Iida,K., Uesugi,H., Sugiyama,Y. and  
Nomoto,H.  
Expressed genes in C.elegans  
Unpublished (1999)  
JOURNAL  
COMMENT On Dec 20, 1995 this sequence version replaced gi:1135120.

CONTACT: Yuji Kohara  
Gene Library Lab  
National Institute of Genetics  
Yata 1111, Mishima, Shizuoka 411, Japan  
Tel: 81-559-81-6854  
Fax: 81-559-81-6855  
Email: ykohara@lab.nig.ac.jp.

FEATURES  
source  
Location/Qualifiers  
1..360

/organism="Caenorhabditis elegans"  
/strain="N2"  
/db\_xref="taxon:6239"  
/clone="yk4946"  
/clone\_lib="Yuji Kohara unpublished cDNA:Strain N2  
hermaphrodite embryo"  
/sex="hermaphrodite"  
/dev\_stage="embryo"  
BASE COUNT 103 a 90 c 109 g 58 t  
ORIGIN

alignment\_scores:  
Quality: 46.00 Length: 19  
Ratio: 3.067 Gaps: 0  
Percent Similarity: 78.947 Percent Identity: 52.632

alignment\_block:

US-08-653-294-30 x AV186046/rev ..

Align seg 1/1 to reverse of: AV186046 from: 1 to: 360

2 ArgLeuAlaThrArgLeuAsnGluArgGluAsnLeuArgThrAlaLe 18  
|||||  
174 CGTTTGTCTACCGGCTTCGGAGCGACGACGATCCGGCTTCGCT 125

18 uArgTyr 20  
|||||  
124 CCGATT 118

seq\_name: gb\_est33:AI777709

seq\_documentation\_block:

LOCUS AI777709 393 bp mRNA EST 29-JUN-1999  
DEFINITION EST258504 tomato susceptible, Cornell Lycopersicon esculentum cDNA  
clone CLES2H21, mRNA sequence.

ACCESSION AI777709  
VERSION AI777709.1 GI:5275666  
KEYWORDS EST.  
SOURCE tomato.  
ORGANISM Lycopersicon esculentum



**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2000, 04:05:43 ; Search time 133.56 seconds  
(without alignments)  
3.547 Million cell updates/sec

Title: US-08-653-294-31

Perfect score: 98

Sequence: 1 YRLAIRLNERYLRLAIRLNER 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 188963 seqs, 23686106 residues

Total number of hits satisfying chosen parameters: 188963

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : A\_Geneseq\_36.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	98	100.0	20	1 R92911	HLA-B2702 CTL modu
2	98	100.0	20	1 W33779	Immunomodulating d
3	54.5	55.6	20	1 R92907	HLA-B2702 CTL modu
4	54.5	55.6	20	1 R95428	HLA-B2702 84-75-84
5	54.5	55.6	20	1 W33778	Immunomodulating d
6	51.5	52.6	20	1 R92909	HLA-B2702 CTL modu
7	51.5	52.6	20	1 W33792	Peptide B2702.84-7
8	49.5	50.5	20	1 R92908	HLA-B2702 CTL modu
9	49.5	50.5	20	1 R95430	Peptide B2702.84-7
10	49.5	50.5	20	1 W33791	Immunomodulatory p
11	49.5	50.5	20	1 W47266	Immunomodulatory p
12	46.5	47.4	20	1 R92910	Peptide B2702.84-7
13	46.5	47.4	20	1 W33793	Immunomodulatory p
14	44.4	44.9	10	1 W47268	Immunomodulatory p
15	44.4	44.9	10	1 W47270	Immunomodulatory p
16	39.3	39.8	10	1 W47272	Immunomodulatory p
17	38.8	38.8	65	1 W20748	H. pylori cytoplas
18	38.8	38.8	272	1 W38650	S. pneumoniae alan
19	38.8	38.8	616	1 R65967	MS2 (male sterilit
20	38.8	38.8	672	1 P71200	Thermotoga beta
21	38.8	38.8	672	1 P60219	Sequence of thermo
22	38.8	38.8	706	1 W31199	Bacillus popilliae
23	38.8	38.8	872	1 W27658	Streptococcus pneu
24	38.8	38.8	3588	1 R34712	Bacillus subtilis
25	37.8	37.8	85	1 W60182	B. thuringiensis t
26	37.8	37.8	485	1 R20796	EHV-4 gc. Nucleic
27	37.8	37.8	597	1 R30168	Novel intestinal o
28	37.8	37.8	599	1 R72319	Laccase Rslac2 pro
29	37.8	37.8	599	1 W60879	Rhizoctonia solani
30	37.8	37.8	599	1 W76311	Rhizoctonia solani
31	36.5	37.2	149	1 W62901	Mutant of the first
32	36.5	37.2	149	1 W51846	Amino acid sequenc
33	36.5	37.2	247	1 W88359	Human lymphocyte a
34	36.5	37.2	338	1 W88361	Human lymphocyte a

35	36.5	37.2	380	1	R13272	Polyhedrin-soluble
36	36.5	37.2	422	1	W88360	Human lymphocyte a
37	36.5	37.2	498	1	R13270	Lymphocyte Activat
38	36.5	37.2	498	1	R87089	Human immunoglobul
39	36	36.7	243	1	P70483	Sequence encoded b
40	36	36.7	355	1	R77034	Rat brain-enriched
41	36	36.7	375	1	R95932	PDGF receptor beta
42	36	36.7	375	1	R96129	PR/TS protein, PRL
43	36	36.7	404	1	R94463	Creatinine amidinoh
44	36	36.7	404	1	W11861	Creatinine amidinoh
45	36	36.7	404	1	W22893	Creatine amidinoh

#### ALIGNMENTS

##### RESULT 1

R92911  
ID R92911 standard; peptide; 20 AA.  
AC R92911:  
DF 16-MAY-1996 (first entry)  
DE HLA-B2702 CTL modulating peptide (B2702.84-75/84-75).  
KW Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;  
KW immunosuppressant; graft versus host disorder; transplantation; therapy;  
KW class I MHC; HLA-B2702.  
OS Synthetic.  
PN WO9526979-A1.  
PD 12-OCT-1995.  
PF 05-APR-1995; U04349.  
PR 05-APR-1994; US-222851.  
PA (STRD) UNIV LELAND STANFORD JUNIOR.  
PI Clayberger C, Krensky AM, Parham P;  
DI WPI: 95-359582/46  
PT Extension of acceptance period of transplants from MHC unmatched  
PT donor hosts - using Class I B75-84 MHC antigen of the recipient  
PT host  
PS Example 15; Page 36; 80pp; English.  
CC R83061-R83095, R83090-R83096 and R92907-R92914 represent fragments of  
CC class I major histocompatibility complex (MHC) antigens. This sequence  
CC is a dimer of residues 75-84 of the alpha-1 domain of the class I MHC  
CC HLA-B2702. These sequences can be used to extend the period of  
CC acceptance by a recipient of a transplant from an MHC unmatched donor.  
CC The peptides are administered to a patient in conjunction with a  
CC subtherapeutic amount of an immunosuppressant. This is administered to  
CC the patient for a limited period of time (compared to the lifetime  
CC administration for current treatments). The peptides particularly  
CC modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)  
CC of the patient.  
SQ Sequence 20 AA;

Query Match 100.0%; Score 98; DB 1; Length 20;  
Best Local Similarity 100.0%; Pred. No. 3.7e-10;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLAIRLNERYLRLAIRLNER 20  
DB 1 YRLAIRLNERYLRLAIRLNER 20

##### RESULT 2

W33779  
ID W33779 standard; peptide; 20 AA.  
AC W33779:  
DF 19-JUN-1998 (first entry)  
DE Immunomodulating dimer peptide #2.  
KW Immunomodulating dimer; immunosuppressant drug; CTL activation;  
KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;  
KW rejection.  
OS Synthetic.  
OS Homo sapiens.  
PN WO9744351-A1.  
PD 27-NOV-1997.

PF 22-MAY-1997; U08689.  
 PR 24-MAY-1996; US-653294.  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI Beulow R, Clayberger C, Krensky AM;  
 DR WPI: 98-086530/08.  
 PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B  
 PT alpha-1 domain, used for preventing rejection of transplants or  
 PT treating autoimmune diseases  
 PS Claim 16; Page 35; 41pp; English.  
 CC This sequence represents a specifically claimed immunomodulating  
 CC dimer peptide of the invention. A peptide-type compound or variant is  
 CC claimed which has immunomodulating activity, including the N-terminal  
 CC acylated and/or C-terminal amidated or esterified forms of up to 60  
 CC amino acids, where the peptide-type compound comprises the formula: A-B,  
 CC where A, B = (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or  
 CC V; aa77 = D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a  
 CC hydrophobic or small amino acid; aa82 = R or L; aa83 = G or R; and aa  
 CC represents amino acid. The sequence in the brackets may optionally be  
 CC absent or truncated at any peptide type bond within the brackets. The  
 CC compounds comprise amino acid sequences related to a Class I HLA-B  
 CC alpha domain (positions 79-84). They can be used to inhibit cytotoxic  
 CC T-lymphocytes (CTL) from undesirably attacking cells in a host or in  
 CC vitro. They can also be used in combination with antigenic peptides or  
 CC proteins of interest to activate CTLs. They can also inhibit the  
 CC proliferation of T cells in response to anti-CD3. The peptide can be  
 CC used for preventing rejection of transplants or for treating autoimmune  
 CC diseases, e.g. diabetes, rheumatoid arthritis and lupus erythematosus.  
 CC The products can also be used for detection and diagnosis.  
 SQ Sequence 20 AA;

Query Match 100.0%; Score 98; DB 1; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 3.7e-10;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLAIRLNERYLRLAIRNER 20  
 |||||  
 DB 1 YRLAIRLNERYLRLAIRNER 20

## RESULT 3

R92907  
 ID R92907 standard; peptide; 20 AA.  
 AC R92907;  
 DT 16-MAY-1996 (first entry)  
 DE HLA-B2702 CTL modulating peptide (B2702.84-75/75-84).  
 KW Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;  
 KW immunosuppressant; graft versus host disorder; transplantation; therapy;  
 KW class I MHC; HLA-B2702.  
 OS Synthetic.  
 PN W09528979-A1.  
 PD 12-OCT-1995.  
 PF 05-APR-1995; U04349.  
 PR 05-APR-1994; US-222851.  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI Clayberger C, Krensky AM, Farham P;  
 DR WPI: 95-358582/46.  
 PT Extension of acceptance period of transplants from MHC unmatched  
 PT donor hosts - using Class I B75-84 MHC antigen of the recipient  
 PT host

PS Example 15; Page 36; 80pp; English.  
 CC R83061-R83085, R83090-R83096 and R92907-R92914 represent fragments of  
 CC Class I major histocompatibility complex (MHC) antigens. This sequence  
 CC is an inverted dimer of residues 75-84 of the alpha-1 domain of the class  
 CC I MHC HLA-B2702. These sequences can be used to extend the period of  
 CC acceptance by a recipient of a transplant from an MHC unmatched donor.  
 CC The peptides are administered to a patient in conjunction with a  
 CC subtherapeutic amount of an immunosuppressant. This is administered to  
 CC the patient for a limited period of time (compared to the lifetime  
 CC administration for current treatments). The peptides particularly  
 CC modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)  
 CC of the patient.  
 SQ Sequence 20 AA;

Query Match 55.6%; Score 54.5; DB 1; Length 20;  
 Best Local Similarity 68.4%; Pred. No. 0.0047;  
 Matches 13; Conservative 2; Mismatches 1; Indels 3; Gaps 1;

QY 1 YRLAIRLNER---YRLAIR 16  
 |||||  
 DB 1 YRLAIRLNERERLRLAIR 19

## RESULT 4

R95428  
 ID R95428 standard; peptide; 20 AA.  
 AC R95428;  
 DT 12-NOV-1996 (first entry)  
 DE HLA-B2702 84-75-84 palindromic.  
 KW HLA; p74; alpha1-helix; human-leucocyte-associated antigen; inhibitor;  
 KW T-cell lysate; membrane protein; mammal; heat shock protein; Hsc70; APC;  
 KW B cell; calcium influx; cytotoxic T lymphocyte; CTL; differentiation;  
 KW cytolysis; antigen presenting cell.  
 OS Synthetic.  
 PN W09513288-A1.  
 PD 18-MAY-1995.  
 PF 10-NOV-1994; U12985.  
 PR 10-NOV-1993; US-150493.  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI Clayberger C, Krensky AM;  
 DR WPI: 95-194027/25.  
 PT Compns. comprising lymphoid surface membrane proteins - which may  
 PT inhibit cytolytic activity and differentiation of CTLs.  
 PS Example; Page 12; 29pp; English.  
 CC R95413, and R95415-R95431 represent palindromes and fragments of  
 CC human-leucocyte-associated antigens. This sequence represents the  
 CC HLA-B2702 84-75-84 palindromic. These sequences can be used to isolate  
 CC the protein p74 from a T-cell lysate. p74 is a T-cell surface membrane  
 CC protein associated with T-cell activation in mammalian T-cells, and is  
 CC also immunologically cross reactive with the heat shock protein Hsc70.  
 CC p74 is found in a limited number of cell types, but is particularly  
 CC expressed on B and T cells. p74 can be isolated by lysis of a suitable  
 CC cell with an amphoteric detergent, and then passed through an affinity  
 CC column containing a covalently bound HLA-B2702 palindromic peptide.  
 CC Compositions comprising the extracellular fragment of p74 combined with  
 CC HLA-B2702.60-84 (see R95416), induces calcium influx, and inhibits  
 CC cytotoxic T lymphocyte (CTL) differentiation or cytolysis. Candidate  
 CC compounds can be screened for their effect on the cytolytic activity of  
 CC T-cells, by combining them with the extracellular portion of p74 and  
 CC determining the amount of binding between the candidate compound and p74.  
 CC Modulation of CTL activity can be inhibited in a cellular composition  
 CC containing T-cells and antigen presenting cells (APCs), by adding to the  
 CC mix the extracellular portion of p74, in an amount sufficient to compete  
 CC with p74 for the binding of the p74 ligand.  
 SQ Sequence 20 AA;

Query Match 55.6%; Score 54.5; DB 1; Length 20;  
 Best Local Similarity 68.4%; Pred. No. 0.0047;  
 Matches 13; Conservative 2; Mismatches 1; Indels 3; Gaps 1;

QY 1 YRLAIRLNER---YRLAIR 16  
 |||||  
 DB 1 YRLAIRLNERERLRLAIR 19

## RESULT 5

W33778  
 ID W33778 standard; peptide; 20 AA.  
 AC W33778;  
 DT 19-JUN-1998 (first entry)  
 DE Immunomodulating dimer peptide #1.  
 KW Immunomodulating dimer; immunosuppressant drug; CTL activation;  
 KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;  
 KW rejection.

OS Synthetic.  
 OS Homo sapiens.  
 PN W09744351-A1.  
 PD 27-NOV-1997.  
 PF 22-MAY-1997; U08689.  
 PR 24-MAY-1996; US-653294.  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI Beulow R, Clayberger C, Krensky AM;  
 DR WPI: 98-086530/08.  
 PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B  
 PT alpha-1 domain, used for preventing rejection of transplants or  
 PT treating autoimmune diseases  
 PS Claim 16; Page 35; 41pp; English.  
 CC This sequence represents a specifically claimed immunomodulating  
 CC dimer peptide of the invention. A peptide-type compound or variant is  
 CC claimed which has immunomodulating activity, including the N-terminal  
 CC acylated and/or C-terminal amidated or esterified forms of up to 60  
 CC amino acids, where the peptide-type compound comprises the formula: A-B,  
 CC where A, B = (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or  
 CC V; aa77 = D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a  
 CC hydrophobic or small amino acid; aa82 = R or L; aa83 = G or R; and aa  
 CC represents amino acid. The sequence in the brackets may optionally be  
 CC absent or truncated at any peptide type bond within the brackets. The  
 CC compounds comprise amino acid sequences related to a Class I HLA-B  
 CC alphas domain (positions 79-84). They can be used to inhibit cytotoxic  
 CC T-lymphocytes (CTL) from undesirably attacking cells in a host or in  
 CC vitro. They can also be used in combination with antigenic peptides or  
 CC proteins of interest to activate CTLs. They can also inhibit the  
 CC proliferation of T cells in response to anti-CD3. The peptide can be  
 CC used for preventing rejection of transplants or for treating autoimmune  
 CC diseases, e.g. diabetes, rheumatoid arthritis and lupus erythematosus.  
 CC The products can also be used for detection and diagnosis.  
 CC Sequence 20 AA;

Query Match 55.6%; Score 54.5; DB 1; Length 20;  
 Best Local Similarity 68.4%; Pred. No. 0.0047;  
 Matches 13; Conservative 1; Mismatches 2; Indels 3; Gaps 1;  
 QY 1 YRLAIRLNER---YRLAIR 16  
 DB 1 YRLAIRLNERRENRLALR 19  
 RESULT 6  
 R92909  
 ID R92909 standard; peptide; 20 AA.  
 AC R92909;  
 DE HLA-B2702 CTL modulating peptide (B2702.84-75/75-84(T)).  
 KW Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;  
 KW immunosuppressant; graft versus host disorder; transplantation; therapy;  
 KW class I MHC; HLA-B2702.  
 OS Synthetic.  
 PN W09526979-A1.  
 PD 12-OCT-1995.  
 PF 05-APR-1995; U04349.  
 PR 05-APR-1994; US-222851.  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI Clayberger C, Krensky AM, Parham P;  
 DR WPI: 95-358582/46.  
 PT Extension of acceptance period of transplants from MHC unmatched  
 PT donor hosts - using Class I B75-84 MHC antigen of the recipient  
 PT host  
 PS Example 15; Page 36; 80pp; English.  
 CC R83061-R83085, R83090-R83096 and R92907-R92913 represent fragments of  
 CC class I major histocompatibility complex (MHC) antigens. This sequence  
 CC is an inverted dimer of residues 75-84 of the alpha-1 domain of the class  
 CC I MHC HLA-B2702. These sequences can be used to extend the period of  
 CC acceptance by a recipient of a transplant from an MHC unmatched donor.  
 CC The peptides are administered to a patient in conjunction with a  
 CC subtherapeutic amount of an immunosuppressant. This is administered to  
 CC the patient for a limited period of time (compared to the lifetime

CC administration for current treatments). The peptides particularly  
 CC modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)  
 CC of the patient.  
 SQ Sequence 20 AA;  
 Query Match 52.6%; Score 51.5; DB 1; Length 20;  
 Best Local Similarity 68.4%; Pred. No. 0.015;  
 Matches 13; Conservative 1; Mismatches 2; Indels 3; Gaps 1;  
 QY 1 YRLAIRLNER---YRLAIR 16  
 DB 1 YRLAIRLNERRENRLALR 19  
 RESULT 7  
 W33792  
 ID W33792 standard; peptide; 20 AA.  
 AC W33792;  
 DE Peptide B2702.84-75/75-84T tested for immunomodulating activity.  
 DE 19-JUN-1998 (first entry)  
 KW Immunomodulating dimer; immunosuppressant drug; CTL activation;  
 KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;  
 KW rejection.  
 OS Synthetic.  
 OS Homo sapiens.  
 PN W09744351-A1.  
 PD 27-NOV-1997.  
 PF 22-MAY-1997; U08689.  
 PR 24-MAY-1996; US-653294.  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI Beulow R, Clayberger C, Krensky AM;  
 DR WPI: 98-086530/08.  
 PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B  
 PT alpha-1 domain, used for preventing rejection of transplants or  
 PT treating autoimmune diseases  
 PS Example 1; Page 19; 41pp; English.  
 CC Peptides W3784-98 and W3778-9 were assayed for their immunomodulating  
 CC activity. A peptide-type compound or variant is claimed which has  
 CC immunomodulating activity, including the N-terminal acylated and/or  
 CC C-terminal amidated or esterified forms of up to 60 amino acids, where  
 CC the peptide-type compound comprises the formula: A-B, where A, B =  
 CC (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or V; aa77 =  
 CC D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a hydrophobic or  
 CC small amino acid; aa82 = R or L; aa83 = G or R; and aa represents amino  
 CC acid. The sequence in the brackets may optionally be absent or truncated  
 CC at any peptide type bond within the brackets. The compounds comprise  
 CC amino acid sequences related to a Class I HLA-B alphas domain (positions  
 CC 79-84). They can be used to inhibit cytotoxic T-lymphocytes (CTL) from  
 CC undesirably attacking cells in a host or in vitro. They can also be  
 CC used in combination with antigenic peptides or proteins of interest to  
 CC activate CTLs. They can also inhibit the proliferation of T cells in  
 CC response to anti-CD3. The peptide can be used for preventing rejection  
 CC of transplants or for treating autoimmune diseases, e.g. diabetes,  
 CC rheumatoid arthritis and lupus erythematosus. The products can also be  
 CC used for detection and diagnosis.  
 CC Sequence 20 AA;  
 Query Match 52.6%; Score 51.5; DB 1; Length 20;  
 Best Local Similarity 68.4%; Pred. No. 0.015;  
 Matches 13; Conservative 1; Mismatches 2; Indels 3; Gaps 1;  
 QY 1 YRLAIRLNER---YRLAIR 16  
 DB 1 YRLAIRLNERRENRLALR 19  
 RESULT 8  
 R92908  
 ID R92908 standard; peptide; 20 AA.  
 AC R92908;  
 DT 16-MAY-1996 (first entry)

DE HLA-B2702 CTL modulating peptide (B2702, 84-75(T)/75-84).  
 KW Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;  
 KW immunosuppressant; graft versus host disorder; transplantation; therapy;  
 KW class I MHC; HLA-B2702.  
 OS Synthetic.  
 PN W09526979-A1.  
 PD 12-OCT-1995.  
 PF 05-APR-1995; U04349.  
 PR 03-APR-1994; US-222851.  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI Clayberger C, Krensky AM, Parham P;  
 DR WPI; 95-358582/46.  
 PT Extension of acceptance period of transplants from MHC unmatched  
 PT donor hosts - using Class I B75-84 MHC antigen of the recipient  
 PT host  
 PS Example 15; Page 36; 80pp; English.  
 CC R83061-R83085, R83090-R83096 and R92907-R92914 represent fragments of  
 CC class I major histocompatibility complex (MHC) antigens. This sequence  
 CC is an inverted dimer of residues 75-84 of the alpha-1 domain of the class  
 CC I MHC HLA-B2702. These sequences can be used to extend the period of  
 CC acceptance by a recipient of a transplant from an MHC unmatched donor.  
 CC The peptides are administered to a patient in conjunction with a  
 CC subtherapeutic amount of an immunosuppressant. This is administered to  
 CC the patient for a limited period of time (compared to the lifetime  
 CC administration for current treatments). The peptides particularly  
 CC modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)  
 CC of the patient.  
 SQ Sequence 20 AA;

Query Match 50.5%; Score 49.5; DB 1; Length 20;  
 Best Local Similarity 63.2%; Pred. NO. 0.031;  
 Matches 12; Conservative 2; Mismatches 3; Gaps 1;

QY 1 YRLAIRLNER---YRLAIR 16  
 Db 1 YRLATRLNERENLRALR 19

## RESULT 9

R95430  
 ID R95430 standard; peptide; 20 AA.  
 AC R95430;  
 DT 12-NOV-1996 (first entry)  
 DE HLA-B2702 84-75(T)/75-84T palindromic.  
 KW HLA; p74; alpha-helix; human-leucocyte-associated antigen; inhibitor;  
 KW T-cell lysate; membrane protein; mammal; heat shock protein; Hsc70; APC;  
 KW B cell; calcium influx; cytotoxic T lymphocyte; CTL; differentiation;  
 KW cytotoxic; antigen presenting cell.  
 OS Synthetic.  
 PN W09513288-A1.  
 PD 18-MAY-1995.  
 PF 10-NOV-1994; U12985.  
 PR 10-NOV-1993; US-150493.  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI Clayberger C, Krensky AM;  
 DR WPI; 95-194027/25.  
 PT Compsns. comprising lymphoid surface membrane proteins - which may  
 PT inhibit cytolytic activity and differentiation of CTLs.  
 PS Example; Page 12; 29pp; English.  
 CC R95413, and R95415-R95431 represent palindromes and fragments of  
 CC human-leucocyte-associated antigens. This sequence represents the  
 CC HLA-B2702 84-75(T)/75-84T palindromic. These sequences can be used to  
 CC isolate the protein p74 from a T-cell lysate. p74 is a T-cell surface  
 CC membrane protein associated with T-cell activation in mammalian T-cells,  
 CC and is also immunologically cross reactive with the heat shock protein  
 CC Hsc70. p74 is found in a limited number of cell types, but is  
 CC particularly expressed on B and T cells. p74 can be isolated by lysis of  
 CC a suitable cell with an amphoteric detergent, and then passed through an  
 CC affinity column containing a covalently bound HLA-B2702 palindromic  
 CC peptide. Compositions comprising the extracellular fragment of p74  
 CC combined with HLA-B2702.60-84 (see R95416), induces calcium influx, and  
 CC inhibits cytotoxic T lymphocyte (CTL) differentiation or cytotoxicity.

CC Candidate compounds can be screened for their effect on the cytolytic  
 CC activity of T-cells, by combining them with the extracellular portion of  
 CC p74 and determining the amount of binding between the candidate compound  
 CC and p74. Modulation of CTL activity can be inhibited in a cellular  
 CC composition containing T-cells and antigen presenting cells (APCs), by  
 CC adding to the mix the extracellular portion of p74, in an amount  
 CC sufficient to compete with p74 for the binding of the p74 ligand.  
 SQ Sequence 20 AA;

Query Match 50.5%; Score 49.5; DB 1; Length 20;  
 Best Local Similarity 63.2%; Pred. NO. 0.031;  
 Matches 12; Conservative 2; Mismatches 3; Gaps 1;

QY 1 YRLAIRLNER---YRLAIR 16  
 Db 1 YRLAIRLNERENLRALR 19

## RESULT 10

W33791  
 ID W33791 standard; peptide; 20 AA.  
 AC W33791;  
 DT 19-JUN-1998 (first entry)  
 DE Peptide B2702, 84-75(T)/75-84 tested for immunomodulating activity.  
 KW Immunomodulating dimer; immunosuppressant drug; CTL activation;  
 KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;  
 KW rejection.  
 OS Synthetic.  
 PN Homo sapiens.  
 PN W09744351-A1.  
 PD 27-NOV-1997.  
 PF 22-MAY-1997; U08689.  
 PR 24-MAY-1996; US-653294.  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI Beulow R, Clayberger C, Krensky AM;  
 DR WPI; 98-086530/08.  
 PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B  
 PT alpha-1 domain, used for preventing rejection of transplants or  
 PT treating autoimmune diseases  
 PS Example 1; Page 19; 41pp; English.  
 CC Peptides W33784-98 and W33778-9 were assayed for their immunomodulating  
 CC activity. A peptide-type compound or variant is claimed which has  
 CC immunomodulating activity, including the N-terminal acylated and/or  
 CC C-terminal amidated or esterified forms of up to 60 amino acids, where  
 CC the peptide-type compound comprises the formula: A-B, where A =  
 CC (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or V; aa77 =  
 CC D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a hydrophobic or  
 CC small amino acid; aa82 = R or L; aa83 = G or R; and aa represents amino  
 CC acid. The sequence in the brackets may optionally be absent or truncated  
 CC at any peptide type bond within the brackets. The compounds comprise  
 CC amino acid sequences related to a Class I HLA-B alpha-1 domain (positions  
 CC 79-84). They can be used to inhibit cytotoxic T-lymphocytes (CTL) from  
 CC undesirably attacking cells in a host or in vitro. They can also be  
 CC used in combination with antigenic peptides or proteins of interest to  
 CC activate CTLs. They can also inhibit the proliferation of T cells in  
 CC response to anti-CD3. The peptide can be used for preventing rejection  
 CC of transplants or for treating autoimmune diseases, e.g. diabetes, be  
 CC rheumatoid arthritis and lupus erythematosus. The products can also be  
 CC used for detection and diagnosis.  
 SQ Sequence 20 AA;

Query Match 50.5%; Score 49.5; DB 1; Length 20;  
 Best Local Similarity 63.2%; Pred. NO. 0.031;  
 Matches 12; Conservative 2; Mismatches 3; Gaps 1;

QY 1 YRLAIRLNER---YRLAIR 16  
 Db 1 YRLATRLNERENLRALR 19

## RESULT 11





DE Immunomodulatory peptide.  
KW Immunomodulator; Class I HLA-B alpha-1 domain; inhibition;  
transplant rejection; treatment; autoimmune disease.  
OS Homo sapiens.  
FH Key  
FT Location/Qualifiers  
FT Misc\_difference 1..10 /note= "at least one of the amino acids is the  
D-isomer  
PN WO9744052-A1.  
PD 27-NOV-1997.  
PF 23-APR-1997; U06705.  
PR 22-MAY-1996; US-651650.  
PA (STRD) UNIV LELAND STANFORD JUNIOR.  
PI Clayberger C, Krensky AM;  
DR WPI: 98-018220/02.  
PT Novel immunomodulatory peptide-type compound - useful for inhibiting  
transplant rejection  
PS Claim 10; Page 36; 41pp; English.  
CC The present sequence is an immunomodulatory peptide, which  
comprises a Class I HLA-B alpha-1 domain sequence. It can be used  
in a pharmaceutical composition together with a subtherapeutic dose  
of an immunosuppressant, to extend the period of acceptance of a  
transplant from a major histocompatibility complex (MHC) unmatched  
donor, i.e. to inhibit transplant rejection. It can also be used in  
the treatment of autoimmune diseases.  
CC Peptides using the D-form amino acids are more effective  
CC Immunomodulators than their diastereomers or enantiomers.  
SQ Sequence 10 AA;

Query Match 44.9%; Score 44; DB 1; Length 10;  
Best Local Similarity 90.0%; Pred. No. 0.11;  
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 YRLAIRLNER 10  
||| |||||  
DB 1 YRLAIRLNER 10

## RESULT 15

W47270  
ID W47270 standard; peptide; 10 AA.  
AC W47270:  
DT 22-MAY-1998 (first entry)  
DE Immunomodulatory peptide.  
KW Immunomodulator; Class I HLA-B alpha-1 domain; inhibition;  
transplant rejection; treatment; autoimmune disease.  
OS Homo sapiens.  
FH Key  
FT Location/Qualifiers  
FT Misc\_difference 1..10 /note= "at least one of the amino acids is the  
D-isomer  
PN WO9744052-A1.  
PD 27-NOV-1997.  
PF 23-APR-1997; U06705.  
PR 22-MAY-1996; US-651650.  
PA (STRD) UNIV LELAND STANFORD JUNIOR.  
PI Clayberger C, Krensky AM;  
DR WPI: 98-018220/02.  
PT Novel immunomodulatory peptide-type compound - useful for inhibiting  
transplant rejection  
PS Claim 10; Page 36; 41pp; English.  
CC The present sequence is an immunomodulatory peptide, which  
comprises a Class I HLA-B alpha-1 domain sequence. It can be used  
in a pharmaceutical composition together with a subtherapeutic dose  
of an immunosuppressant, to extend the period of acceptance of a  
transplant from a major histocompatibility complex (MHC) unmatched  
donor, i.e. to inhibit transplant rejection. It can also be used in  
the treatment of autoimmune diseases.  
CC Peptides using the D-form amino acids are more effective  
CC Immunomodulators than their diastereomers or enantiomers.

SQ Sequence 10 AA;

Query Match 44.9%; Score 44; DB 1; Length 10;  
Best Local Similarity 90.0%; Pred. No. 0.11;  
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLAIRLNER 10  
||| |||||  
DB 1 YRLAIRLNER 10

Search completed: February 8, 2000, 04:05:44  
Job time: 9361 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 7, 2000, 18:04:39 ; Search time 111.22 seconds  
(without alignments)  
8.482 Million cell updates/sec

Title: US-08-653-294-31

Perfect score: 98

Sequence: 1 YRLAIRLNEYRLAIRLNER 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 142080 seqs, 47169319 residues

Total number of hits satisfying chosen parameters: 142080

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

PIR\_62:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	42.5	43.4	414	2	T15947	hypothetical prote
2	42	42.9	846	2	A30889	integrin beta chai
3	41	41.8	223	2	F70685	probable phosphol
4	41	41.8	253	2	JS0595	rod-core linker po
5	41	41.8	254	2	S23475	rod-core linker po
6	41	41.8	289	2	A69405	hydroxymethylbilan
7	41	41.8	361	1	S34613	5-exo-hydroxycamph
8	41	41.8	413	2	F70211	hypothetical prote
9	41	41.8	584	2	S06696	hypothetical prote
10	40	40.8	192	2	H70375	conserved hypothet
11	40	40.8	388	2	S15593	hypothetical prote
12	40	40.8	473	2	A48949	beta-glucosidase,
13	40	40.8	526	2	S75812	hypothetical prote
14	40	40.8	579	2	A72367	oligopeptide ABC t
15	40	40.8	1502	1	RGB1H1	CyC1/CIP3 transcri
16	40	40.8	2208	2	A37860	calcium channel pr
17	40	40.8	2352	2	T05077	splicing factor-li
18	39.5	40.3	540	2	S76869	hypothetical prote
19	39.5	40.3	1110	1	A70652	probable serine/th
20	39	39.8	144	2	A21047	ribosomal mobile e
21	39	39.8	320	2	G70177	hypothetical prote
22	39	39.8	717	2	S31035	retrovirus-related
23	39	39.8	717	2	S31034	retrovirus-related
24	39	39.8	868	1	S45757	replication licens
25	39	39.8	2077	1	W2B24	240k tegument prot
26	39	39.8	2078	2	T09326	tegument protein -
27	38.5	39.3	193	2	H72776	hypothetical prote
28	38.5	39.3	349	1	R0ECGL	nitrogen regulatio
29	38	38.8	137	2	S43885	3-isopropylmalate
30	38	38.8	279	1	S32869	secretion protein

31	38	38.8	289	2	C72347	motility protein B
32	38	38.8	346	2	G64182	ADP-heptose--LPS h
33	38	38.8	372	2	S46344	env polyprotein -
34	38	38.8	387	2	D89752	two-component sens
35	38	38.8	393	2	JC6176	tumor suppressor p
36	38	38.8	395	2	T01392	leucine-rich repea
37	38	38.8	446	2	S59646	hypothetical prote
38	38	38.8	447	2	A72358	conserved hypothet
39	38	38.8	447	2	T04506	hypothetical prote
40	38	38.8	596	2	S33804	male sterility pro
41	38	38.8	616	2	T08096	beta-galactosidase
42	38	38.8	672	1	A29836	internal virion pr
43	38	38.8	747	1	HIBPC7	hypothetical prote
44	38	38.8	1170	2	A72287	surfactin syntheta
45	38	38.8	3588	2	I40485	

#### ALIGNMENTS

##### RESULT 1

T15947

hypothetical protein F01f1.4 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999

C:Accession: T15947

R:Miller, N.

submitted to the EMBL Data Library, July 1995

A:Description: The sequence of C. elegans cosmid F01f1.

A:Reference number: Z18435

A:Accession: T15947

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-414 <MIL>

A:Cross-references: EMBL:U13070; NID:G529697; PID:G529701; PIDN:AAC46638.1; CESP:F01F

A:Experimental source: strain Bristol N2

C:Genetics:

A:Gene: CESP:F01F1.4

A:Introns: 17/3; 51/3; 120/3; 365/2

Query Match 43.4%; Score 42.5; DB 2; Length 414;  
Best Local Similarity 52.6%; Pred. No. 23;  
Matches 10; Conservative 4; Mismatches 4; Indels 1; Gaps 1;

QY 2 RLAIRLNEYRLAIRLNER 20

DB 62 KAAMELNERYQLDLS-NER 79

##### RESULT 2

A30889

integrin beta chain precursor - fruit fly (Drosophila melanogaster)

C:Species: Drosophila melanogaster

C>Date: 01-Dec-1989 #sequence\_revision 01-Dec-1989 #text\_change 20-Aug-1999

C:Accession: A30889

R:MacKrell, A.J.; Blumberg, B.; Haynes, S.R.; Fessler, J.H.

Proc. Natl. Acad. Sci. U.S.A. 85, 2633-2637, 1988

A:Title: The lethal myospheroid gene of Drosophila encodes a membrane protein homolog

A:Reference number: A30889; MUID:88190122

A:Accession: A30889

A:Molecule type: mRNA

A:Residues: 1-846 <MAC>

A:Cross-references: GB:J03251; NID:G157954; PIDN:AAA28714.1; PID:G157955

C:Genetics:

A:Gene: FlyBase:mys

A:Cross-references: FlyBase:FBgn0004657

C:Superfamily: integrin beta chain; laminin-type EGF-like homology

C:Keywords: cell adhesion; cytoskeleton; transmembrane protein

Query Match 42.9%; Score 42; DB 2; Length 846;  
Best Local Similarity 38.9%; Pred. No. 58;

Matches 7; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

Qy 2 RLALRLNRYRLAIRLNE 19  
 |||:||||:| :| :|  
 Db 162 RLALRVNEKHNIKISYSQ 179

RESULT 3  
 F70685  
 probable phosphoglycerate mutase 2 - Mycobacterium tuberculosis (strain H37RV)  
 C:Species: Mycobacterium tuberculosis  
 C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 17-Jul-1998  
 C:Accession: F70685  
 R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
 Nature 393, 537-544, 1998  
 A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
 A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
 A:Reference number: A70500; MUID:98295987  
 A:Accession: F70685  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-223 <COL>  
 A:Cross-references: GB:881368; GB:AL123456; NID:g3261656; PID:e1299808; PID:g3261659  
 A:Experimental source: strain H37RV  
 C:Genetics:  
 A:Gene: RV2419C

Query Match 41.8%; Score 41; DB 2; Length 223;  
 Best Local Similarity 40.0%; Pred. No. 21;  
 Matches 8; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

Qy 1 YRLAIRLNERYLRLNER 20  
 | | : | | | : | : | : |  
 Db 62 YDTAVKLGERTGLVRVDTR 81

RESULT 4  
 JS0595  
 rod-core linker polypeptide cpcG4 - Anabaena sp. (strain PCC 7120)  
 C:Species: Anabaena sp.  
 C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 24-Oct-1997  
 C:Accession: JS0595; PS0243  
 R:Bryant, D.A.; Stirewalt, V.L.; Glauser, M.; Frank, G.; Sidler, W.; Zuber, H.  
 Gene 107, 91-99, 1991  
 A:Title: A small multigene family encodes the rod-core linker polypeptides of Anabaena  
 A:Reference number: JS0592; MUID:92077441  
 A:Accession: JS0595  
 A:Molecule type: DNA  
 A:Residues: 1-253 <GRY>  
 A:Cross-references: GB:M80435; NID:g142097; PID:g142102  
 A:Accession: PS0243  
 A:Molecule type: protein  
 A:Residues: 2-23 <GRV1>  
 A:Comment: Linker polypeptides determine the positions of phycobilliproteins within the P  
 C:Genetics:  
 A:Gene: cpcG4  
 F:2-253/Product: rod-core linker polypeptide cpcG4 #status experimental <MAT>

Query Match 41.8%; Score 41; DB 2; Length 253;  
 Best Local Similarity 53.8%; Pred. No. 23;  
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 YRLAIRLNERYLRL 13  
 ||| : | : | : |  
 Db 95 YRLVSVNNRYRL 107

RESULT 5  
 S23475

rod-core linker polypeptide cpcG3 - Fischerella sp.  
 C:Species: Fischerella sp.  
 C:Date: 07-Apr-1994 #sequence\_revision 07-Apr-1994 #text\_change 07-May-1999  
 C:Accession: S23475; S16060  
 R:Glauser, M.; Stirewalt, V.L.; Bryant, D.A.; Sidler, W.; Zuber, H.  
 Eur. J. Biochem. 205, 927-937, 1992  
 A:Title: Structure of the genes encoding the rod-core linker polypeptides of Mastigoc  
 A:Reference number: S23472; MUID:92249337  
 A:Accession: S23475  
 A:Molecule type: DNA  
 A:Residues: 1-254 <GLA>  
 A:Cross-references: EMBL:X59763; NID:g44397; PID:g44401  
 A:Note: the source is designated as Mastigocladus laminosus  
 C:Genetics:  
 A:Gene: cpcG3

Query Match 41.8%; Score 41; DB 2; Length 254;  
 Best Local Similarity 53.8%; Pred. No. 24;  
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 YRLAIRLNERYLRL 13  
 ||| : | : | : |  
 Db 95 YRLVSVNNRYRL 107

RESULT 6  
 A69405  
 hydroxymethylbilane synthase (EC 4.3.1.18) - Archaeoglobus fulgidus  
 N:Alternate names: porphobilinogen deaminase  
 C:Species: Archaeoglobus fulgidus  
 C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 16-Jul-1999  
 C:Accession: A69405  
 R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dod  
 :; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E  
 Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.  
 Nature 390, 364-370, 1997  
 A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiaach, P.; Kaine, B.P.; Sykes,  
 Smith, H.O.; Woese, C.R.; Venter, J.C.  
 A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing arch  
 A:Reference number: A69250; MUID:98049343  
 A:Accession: A69405  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-289 <KLE>  
 A:Cross-references: GB:AE001018; GB:AE000782; NID:g2689341; PID:AA90000.1; PID:g264  
 C:Superfamily: hydroxymethylbilane synthase  
 C:Keywords: ammonia-lyase; carbon-nitrogen lyase; porphyrin biosynthesis  
 F:234/Modified site: dipyrrolylmethanemethyl (Cys) (covalent) #status predicted

Query Match 41.8%; Score 41; DB 2; Length 289;  
 Best Local Similarity 50.0%; Pred. No. 27;  
 Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 2 RLAIRLNERYLRLAIRL 17  
 : | | | | : | : | : | : | : |  
 Db 19 KVAERLKERIEVEIRI 34

RESULT 7  
 S34613  
 5-exo-hydroxycamphor dehydrogenase (EC 1.1.1.-) - Pseudomonas putida plasmid CAM  
 C:Species: Pseudomonas putida  
 C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
 C:Accession: S34613; A29844  
 R:Aramaki, H.; Koga, H.; Sagara, Y.; Hosoi, M.; Horiuchi, T.  
 Biochim. Biophys. Acta 1174, 91-94, 1993  
 A:Title: Complete nucleotide sequence of the 5-exo-hydroxycamphor dehydrogenase gene  
 A:Reference number: S34613; MUID:93326643  
 A:Accession: S34613  
 A:Status: not compared with conceptual translation  
 A:Molecule type: DNA

A:Residues: 1-361 <ARA>  
 A:Cross-references: GB:D14680; GB:D14452; GB:D14453; NID:g473744  
 A:Experimental source: PpG1; ATCC 17453; CAM plasmid  
 R.Koga, H.; Aramaki, H.; Yamaguchi, E.; Takeuchi, K.; Horiuchi, T.; Gunsalus, I.C.  
 J. Bacteriol. 166, 1089-1095, 1986  
 A:Title: camR, a negative regulator locus of the cytochrome P-450-cam hydroxylase operon  
 A:Reference number: A29844; MUID:86223770  
 A:Accession: A29844

A:Molecule type: DNA  
 A:Residues: 1-96, 'RAIAV' <KOG>  
 A:Cross-references: GB:M13471; NID:g151117  
 A:Note: this sequence has been revised in reference S34613  
 C:Genetics:  
 A:Gene: camD  
 A:Genome: plasmid  
 C:Complex: homodimer  
 C:Function:  
 A:Description: catalyzes formation of 2,5-diketo-camphane from 5-exo-hydroxycamphor  
 C:Superfamily: alcohol dehydrogenase; long-chain alcohol dehydrogenase homology  
 C:Keywords: NAD; oxidoreductase; zinc  
 F:25-350/Domain: long-chain alcohol dehydrogenase homology <LADH>  
 F:178-207/Region: beta-alpha-beta NAD nucleotide-binding fold  
 F:40-62,158/Binding site: zinc, catalytic (Cys, His, Cys) #status predicted  
 F:98,101,104,112/Binding site: zinc, noncatalytic (Cys) #status predicted

Query Match 41.8%; Score 41; DB 1; Length 361;  
 Best Local Similarity 47.4%; Pred. No. 34;  
 Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 2 RLAIRLNRYRLAIRNER 20  
 DB 316 QLAARLQDRIPLADLTQR 334

RESULT 8  
 F70211  
 hypothetical protein BBA38 - Lyme disease spirochete plasmid A/lp54  
 C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
 C:Date: 13-Feb-1998 #sequence\_revision 13-Feb-1998 #text\_change 05-Jun-1998  
 C:Accession: F70211  
 R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White  
 son, D.; Peterson, J.; Kervage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,  
 Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.  
 Nature 390, 580-586, 1997  
 A:Authors: Smith, H.O.; Venter, J.C.  
 A:Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.  
 A:Reference number: A70100; MUID:98065943  
 A:Accession: F70211  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-413 <KLE>  
 A:Cross-references: GB:AE000790; NID:g2690224; PID:g2690299; TIGR:BBA38  
 A:Experimental source: strain B31  
 C:Genetics:  
 A:Genome: plasmid

Query Match 41.8%; Score 41; DB 2; Length 413;  
 Best Local Similarity 46.7%; Pred. No. 39;  
 Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 5 IRLNRYRLAIRNE 19  
 DB 351 LKLNRYRLNMFNE 365

RESULT 9  
 S06696  
 hypothetical protein 1 - Zymomonas mobilis plasmid 2M2  
 C:Species: Zymomonas mobilis  
 C:Date: 28-Feb-1990 #sequence\_revision 28-Feb-1990 #text\_change 09-Sep-1997  
 C:Accession: S06696

R.Misawa, N.; Nakamura, K.  
 submitted to the EMBL Data Library, February 1989  
 A:Reference number: S06695  
 A:Accession: S06696  
 A:Molecule type: DNA  
 A:Residues: 1-584 <MIS>  
 A:Cross-references: EMBL:X14438; NID:g48656; PID:g48657  
 C:Genetics:  
 A:Genome: plasmid

Query Match 41.8%; Score 41; DB 2; Length 584;  
 Best Local Similarity 61.5%; Pred. No. 56;  
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 7 LNERVRLAIRNE 19  
 DB 355 LNDLYRSKIRINE 367

RESULT 10  
 H70375  
 conserved hypothetical protein aq\_880 - Aquifex aeolicus  
 C:Species: Aquifex aeolicus  
 C:Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 08-May-1998  
 C:Accession: H70375  
 R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.;

V.  
 Nature 392, 353-358, 1998  
 A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.  
 A:Reference number: A70300; MUID:98196666  
 A:Accession: H70375  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-192 <AQF>  
 A:Cross-references: GB:AE000712; NID:g2983411; PID:g2983420; GB:AE000657  
 A:Experimental source: strain VF5  
 C:Genetics:  
 A:Gene: aq\_880

Query Match 40.8%; Score 40; DB 2; Length 192;  
 Best Local Similarity 53.3%; Pred. No. 25;  
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 RLAIRLNRYRLAIR 16  
 DB 119 KLIALRLREKREAR 133

RESULT 11  
 S15593  
 hypothetical protein (insertion sequence ISH27-3) - Halobacterium halobium  
 C:Species: Halobacterium halobium  
 C:Date: 19-Mar-1997 #sequence\_revision 25-Apr-1997 #text\_change 30-Jun-1998  
 C:Accession: S15593  
 R:Pfeiffer, F.; Blaseio, U.  
 Nucleic Acids Res. 18, 6921-6925, 1990  
 A:Title: Transposition burst of the ISH27 insertion element family in Halobacterium h  
 A:Reference number: S15591; MUID:91088266  
 A:Accession: S15593  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-388 <PFE>  
 A:Cross-references: EMBL:X54434  
 A:Note: the authors translated the initiation codon GTG for residue 1 as Val  
 C:Genetics:  
 A:Mobile element: insertion sequence ISH27-3  
 A:Start codon: GTG

Query Match 40.8%; Score 40; DB 2; Length 388;  
 Best Local Similarity 77.8%; Pred. No. 52;

Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 8 NERVRLAIR 16

Db 151 NKRYRLAVR 159

# RESULT 12

A48949

beta-glucosidase, BglB - Microbispora bispora

C:Species: Microbispora bispora

C:Date: 19-Dec-1993 #sequence\_revision 25-Apr-1997 #text\_change 22-Jun-1999

C:Accession: A48949

R:Wright, R.M.; Yablonsky, M.D.; Shalita, Z.P.; Goyal, A.K.; Eveleigh, D.E.

Appl. Environ. Microbiol. 58, 3455-3465, 1992

A:Title: Cloning, characterization, and nucleotide sequence of a gene encoding Microbispora

A:Reference number: A48949; MUID:93128922

A:Contents: NRRL 15568

A:Accession: A48949

A>Status: preliminary

A:Molecule type: nucleic acid

A:Residues: 1-473 <WRI>

A:Cross-references: GB:M97265; NID:gl49827; PIDN:AAA25311.1; PID:gl49828

A:Note: sequence extracted from NCBI backbone (NCBIN:122282, NCBI:P:122284)

C:Superfamily: Agrobacterium beta-glucosidase

# Query Match

Best Local Similarity 40.8%; Score 40; DB 2; Length 473;

Matches 11; Conservative 2; Mismatches 6; Indels 2; Gaps 1;

Qy 1 YRLAI--RLNERYRLAIRLNE 19

Db 176 YALAVHRLGDRVRCWITLNE 196

# RESULT 13

S75812

hypothetical protein slr1270 - Synechocystis sp. (strain PCC 6803)

C:Species: Synechocystis sp.

A:Variety: PCC 6803

C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 21-Aug-1998

C:Accession: S75812

O. K.; Okumura, S.; Shingo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda

DNA Res. 3, 109-136, 1996

A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis

A:Reference number: S74322; MUID:97061201

A:Accession: S75812

A>Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-526 <KAN>

A:Cross-references: EMBL:D90913; GB:AB001339; NID:gl653348; PID:d1019004; PID:gl653357

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

# Query Match

Best Local Similarity 40.8%; Score 40; DB 2; Length 526;

Matches 10; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 2 RLAIRLNERYRLAIRLNER 20

Db 98 RQAITLDEAIDLALNNEQ 116

# RESULT 14

A72367

oligopeptide ABC transporter, periplasmic oligopeptide-binding protein - Thermotoga maritima

C:Species: Thermotoga maritima

C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 11-Jun-1999

C:Accession: A72367

R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey

Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, C.M.

Nature 399, 323-329, 1999

A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome

A:Reference number: A72200; MUID:99287316

A:Accession: A72367

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-579 <ARN>

A:Cross-references: GB:AE001728; GB:AE000512; NID:94981027; PID:94981044; TIGR:TM0531

A:Experimental source: strain MS98

C:Genetics:

A:Gene: TM0531

# Query Match

Best Local Similarity 40.8%; Score 40; DB 2; Length 579;

Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 1 YRLAIRLNERYLRLAIR 16

Db 158 YTVKTLKETRYRLAVR 173

# RESULT 15

RGBYH1

CYC1/CYP3 transcription activator - yeast (Saccharomyces cerevisiae)

N:Alternate names: protein L9672.1; protein YLR255w; regulatory protein CYP1; regulat

C:Species: Saccharomyces cerevisiae

C:Date: 30-Sep-1991 #sequence\_revision 23-Feb-1996 #text\_change 16-Jul-1999

C:Accession: S59400; A31312; S15447; S05804; S15446

R:Johnson, D.

submitted to the EMBL Data Library, February 1995

A:Description: The sequence of S. cerevisiae cosmid 9672.

A:Reference number: S59386

A:Accession: S59400

A:Molecule type: DNA

A:Residues: 1-1502 <JOH>

A:Cross-references: EMBL:U20865; NID:g662330; PIDN:AAB67387.1; PID:g662331; MIPS:YLR2

A:Experimental source: strain S288C (AB972)

R:Preifer, K.; Kim, K.S.; Kogan, S.; Guarente, L.

Cell 56, 291-301, 1989

A:Title: Functional dissection and sequence of yeast HAP1 activator.

A:Reference number: A31312; MUID:89106221

A:Accession: A31312

A:Molecule type: DNA

A:Residues: 1-144, 'I', 146-322, 'R', 324-454, 'N', 456-507, 'M', 509-586, 'K', 588-882, 'N', 884

A:Cross-references: EMBL:J03152; NID:gl71645; PIDN:AAA34662.1; PID:gl71646

R:Creusot, F.; Verdier, J.; Gaisne, M.; Slonimski, P.P.

J. Mol. Biol. 204, 263-276, 1988

A:Title: CYPI (HAP1) regulator of oxygen-dependent gene expression in yeast. I. Overa

A:Reference number: S15447; MUID:89125585

A:Accession: S15447

A:Molecule type: DNA

A:Residues: 1-62, 'R', 64-1305, 'Y', 1306-1470, 'LVDFYRADFPPIWE' <CREI>

A:Cross-references: EMBL:X13793

A:Note: the sequence is from mutant CYP1-18

C:Genetics:

A:Gene: SGD:HAP1; CYP1

A:Cross-references: SGD:S0004246; MIPS:YLR256w

A:Map position: 12R

C:Superfamily: regulatory protein HAP1; GAL4 zinc binuclear cluster homology

C:Keywords: DNA binding; heme binding; transcription regulation; zinc finger

F:1-148/Domain: DNA binding #status predicted <DNA>

F:59-98/Domain: GAL4 zinc binuclear cluster homology <GAL4>

F:64-84/Region: zinc finger CCCC motif

F:177-189/Region: glutamine-rich

F:245-445/Domain: heme binding #status predicted <HEM>

F:299-304, 323-328, 347-378, 389-394, 415-420/Region: 6-residue repeats

F:1308-1481/Domain: activation element #status predicted <ACT>

F:1388-1481/Region: acidic

Query Match 40.8%; Score 40; DB 1; Length 1502;  
Best Local Similarity 38.9%; Pred. No. 2.1e+02;  
Matches 7; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

Qy 2 RLAIRLNERYRLAIRLNE 19  
:| :|:::| |:::|  
Db 1124 KLTQLSKKRYNYAIRMNK 1141

Search completed: February 7, 2000, 18:04:40  
Job time: 22206 sec

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: February 8, 2000, 01:26:00 ; Search time 68.63 seconds  
(without alignments)  
8.703 Million cell updates/sec

Title: US-08-653-294-31

Perfect score: 98

Sequence: 1 YRLAIRLNERYLRLAIRLNER 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 82229 seqs, 29864866 residues

Total number of hits satisfying chosen parameters: 82229

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : SwissProt\_38:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	42.9	846	1	ITBX_DROME
2	41	41.8	252	1	P1584 drosophila
3	41	41.8	253	1	P2989 anabaena sp
4	41	41.8	289	1	P29733 mastigoclad
5	41	41.8	361	1	O29026 archaeoglob
6	41	41.8	584	1	P09347 pseudomonas
7	40	40.8	323	1	P15255 zymomonas m
8	40	40.8	473	1	P51947 xenopus lae
9	40	40.8	1483	1	P38645 microbispor
10	39	39.8	312	1	P12351 saccharomyc
11	39	39.8	381	1	O29537 canis fami
12	39	39.8	868	1	P29469 saccharomyc
13	39	39.8	2077	1	P30002 herpes simp
14	39	39.8	2077	1	P52340 herpes simp
15	38.5	39.3	349	1	P06712 escherichia
16	38	38.8	137	1	P50181 neisseria l
17	38	38.8	279	1	P31712 erwinia car
18	38	38.8	346	1	P45042 haemophilus
19	38	38.8	373	1	P79734 brachydanio
20	38	38.8	387	1	P40758 bacillus su
21	38	38.8	393	1	O09185 cricetus
22	38	38.8	446	1	O09718 schistosacch
23	38	38.8	616	1	O08891 arabidopsis
24	38	38.8	672	1	P39668 bacillus st
25	38	38.8	747	1	P03725 bacterioph
26	38	38.8	3587	1	P27206 bacillus su
27	37.5	38.3	349	1	P41788 salmonella
28	37	37.8	132	1	P03563 tomato gold
29	37	37.8	134	1	VAL3_SLCV
30	37	37.8	172	1	DINB_BACSU
31	37	37.8	179	1	YP1B_BACSU
32	37	37.8	187	1	RL5_MYCTU
33	37	37.8	189	1	YTRLE_LPEPI
34	37	37.8	232	1	CYSH_SYNPT
					Q55309 synechococc

35	37	37.8	287	1	Y030_HAEIN	Q57092 haemophilus
36	37	37.8	327	1	EBGR_ECOLI	P06846 escherichia
37	37	37.8	396	1	P53_MESAU	Q00366 mesocricetu
38	37	37.8	424	1	PQ05_CAEEL	Q09535 caenorhabdl
39	37	37.8	445	1	PHR_MEITH	P12769 methanobact
40	37	37.8	485	1	VGLC_HSV4	P22596 equine herp
41	37	37.8	534	1	CN9A_MOUSE	O70628 mus musculu
42	37	37.8	580	1	GPC3_HUMAN	P51654 homo sapien
43	37	37.8	593	1	CN9A_HUMAN	O76083 homo sapien
44	37	37.8	597	1	GPC3_RAT	P13265 rattus norv
45	37	37.8	599	1	LAC2_THACU	Q02075 thanatephor

## ALIGNMENTS

RESULT 1						
ITBX_DROME						
ID	ITBX_DROME	STANDARD;	PRT;	846 AA.		
AC	P11584;					
DT	01-OCT-1989 (Rel. 12, Created)					
DT	01-OCT-1989 (Rel. 12, Last sequence update)					
DT	15-JUL-1998 (Rel. 36, Last annotation update)					
DE	INTEGRIN BETA-SUBUNIT MYOSPHEROID PRECURSOR.					
GN	L(1)MYS.					
OS	Drosophila melanogaster (Fruit fly).					
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;					
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;					
OC	Ephyroidea; Drosophilidae; Drosophila.					
[1]						
RN	SEQUENCE FROM N.A.					
RP	MEDLINE: 88190122.					
RA	MACKRELL A.J., BLUMBERG B., HAYNES S.R., FESSLER J.H.;					
RT	"The lethal myospheroid gene of Drosophila encodes a membrane protein					
RT	homologous to vertebrate integrin beta subunits.";					
RL	Proc. Natl. Acad. Sci. U.S.A. 85:2633-2637(1988).					
[2]						
RP	CHARACTERIZATION.					
EX	MEDLINE: 94163982.					
RA	GRINBLAT Y., ZUSMAN S., YEE G., HYNES R.O., KAFATOS F.C.;					
RT	"Functions of the cytoplasmic domain of the beta PS integrin subunit					
RT	during Drosophila development.";					
RL	Development 120:91-102(1994).					
CC	-1- FUNCTION: NOT KNOWN. PROBABLY PLAYS A ROLE IN CELL ADHESION.					
CC	-1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.					
CC	-1- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.					
CC						
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration					
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -					
CC	the European Bioinformatics Institute. There are no restrictions on its					
CC	use by non-profit institutions as long as its content is in no way					
CC	modified and this statement is not removed. Usage by and for commercial					
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/					
CC	or send an email to license@isb-sib.ch).					
CC						
DR	EMBL: J03251; AAA28714.1; .					
DR	PIR: A30889; A30889.					
DR	FLYBASE: FBgn0004657; mys.					
DR	PROSITE: PS00243; INTEGRIN_BETA; 3.					
DR	PROSITE: PS00022; EGF_1; UNKNOWN_2.					
DR	PROSITE: PS01186; EGF_2; UNKNOWN_1.					
DR	PFAM: PF00362; Integrin_B; 1.					
KW	Integrin; Cell adhesion; Transmembrane; Glycoprotein; Repeat;					
KW	Extracellular matrix; Cytoskeleton; Signal.					
FT	SIGNAL 1 23					
FT	POTENTIAL.					
FT	CHAIN 24 846					
FT	INTEGRIN BETA-SUBUNIT MYOSPHEROID.					
FT	DOMAIN 24 776					
FT	EXTRACELLULAR (POTENTIAL).					
FT	TRANSMEM 777 799					
FT	POTENTIAL.					
FT	DOMAIN 800 846					
FT	CYTOPLASMIC (POTENTIAL).					
FT	DOMAIN 115 143					
FT	SER-RICH.					
FT	DOMAIN 507 687					
FT	CYSTEINE-RICH REPEATS.					
FT	REPEAT 507 560					
FT	I.					
FT	REPEAT 507 605					
FT	II.					

FT	REPEAT	606	646	III.
FT	REPEAT	647	687	IV.
FT	CARBOHYD	72	72	POTENTIAL.
FT	CARBOHYD	266	266	POTENTIAL.
FT	CARBOHYD	277	277	POTENTIAL.
FT	CARBOHYD	403	403	POTENTIAL.
FT	CARBOHYD	428	428	POTENTIAL.
FT	CARBOHYD	557	557	POTENTIAL.
FT	CARBOHYD	603	603	POTENTIAL.
FT	CARBOHYD	644	644	POTENTIAL.
FT	CARBOHYD	718	718	POTENTIAL.
SQ	SEQUENCE	846 AA;	92687 MW; 9906C2F9 CRC32;	

  

Query Match	42.98;	Score 42;	DB 1;	Length 846;
Best Local Similarity	38.98;	Pred. No. 27;		
Matches	7;	Conservative	7;	Mismatches 4; Indels 0; Gaps 0;

  

QY	2 RLAIRLNERYRLAIRLNE	19
DB	162 RLALRVNEKHNIKISYSQ	179
	:    : : :	

  

RESULT	2
PYG4_ANASP	
ID	PYG4_ANASP STANDARD; PRT; 252 AA.
AC	P29989;
DT	01-APR-1993 (Rel. 25, Created)
DT	01-JUL-1993 (Rel. 26, Last sequence update)
DT	01-FEB-1994 (Rel. 28, Last annotation update)
DE	PHYCIBILISOME ROD-CORE LINKER POLYPEPTIDE CPGC4 (L-RC 29.2).
GN	CPGC4.
OS	Anabaena sp. (strain PCC 7120).
OC	Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Anabaena.
[1]	SEQUENCE FROM N.A., AND SEQUENCE OF 1-22.
RP	MEDLINE; 92077441.
RA	BRYANT D.A., STIREWALT V.L., GLAUSER M., FRANK G., SIDLER W.,
RT	ZUBER H.;
RT	"A small multigene family encodes the rod-core linker polypeptides of
RT	Anabaena sp. PCC7120 phycobilisomes.";
RL	Gene 107:91-99(1991).
CC	-!- FUNCTION: ROD-CORE LINKER PROTEIN REQUIRED FOR ATTACHMENT OF
CC	PHYCOCYANIN TO ALLOPHYCOCYANIN IN CORES OF PHYCIBILISOMES.
CC	-!- FUNCTION: LINKER POLYPEPTIDES DETERMINE THE STATE OF AGGREGATION
CC	AND THE LOCATION OF THE DISC-SHAPED PHYCIBILIPROTEIN UNITS WITHIN
CC	THE PHYCIBILISOME AND MODULATE THEIR SPECTROSCOPIC PROPERTIES IN
CC	ORDER TO MEDIATE A DIRECTED AND OPTIMAL ENERGY TRANSFER.
CC	-!- SUBUNIT: THE PHYCIBILISOME IS A HEMIDISCOIDAL STRUCTURE THAT IS
CC	COMPOSED OF TWO DISTINCT SUBSTRUCTURES: A CORE COMPLEX (THAT
CC	CONTAINS PHYCIBILIPROTEINS) AND A NUMBER OF RODS RADIATING FROM
CC	THE CORE.
CC	-!- SIMILARITY: TO OTHER PHYCIBILISOME LINKER PROTEINS.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation at
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announcement/">http://www.isb-sib.ch/announcement/</a>
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).
CC	-----
CC	EMBL; M80435; AAA22039.1; -.
DR	PIR; J50595; J50595.
DR	PFAM; PF00427; PDB_linker_poly; 1.
DR	Phycobilisome; Photosynthesis; Multigene family.
FT	INIT MET 0
SQ	SEQUENCE 252 AA; 29191 MW; 0C1A6468 CRC32;

  

Query Match	41.88;	Score 41;	DB 1;	Length 252;
Best Local Similarity <td>53.88;</td> <td>Pred. No. 10;</td> <td></td> <td></td>	53.88;	Pred. No. 10;		
Matches	7;	Conservative	2;	Mismatches 4; Indels 0; Gaps 0;

RX MEDLINE: 93325643.  
 RA ARAMAKI H., KOGA H., SAGARA Y., HOSOI M., HORIUCHI T.:  
 RT "complete nucleotide sequence of the 5-exo-hydroxycamphor  
 dehydrogenase gene on the CAM plasmid of Pseudomonas putida (ATCC  
 RT 17453).";  
 RL Biochim. Biophys. Acta 1174:91-94(1993).  
 RN [2]  
 RP SEQUENCE OF 1-100 FROM N.A., AND SEQUENCE OF 1-45.  
 RC STRAIN-ATCC 17453;  
 RX MEDLINE: 86223770.  
 RA KOGA H., ARAMAKI H., YAMAGUCHI E., TAKEUCHI K., HORIUCHI T.,  
 RA GUNSAIUS I.C.:  
 RT "camk, a negative regulator locus of the cytochrome P-450cam  
 hydroxylase operon.";  
 RL J. Bacteriol. 166:1089-1095(1986).  
 CC -1- CATALYTIC ACTIVITY: 5-EXO-HYDROXYCAMPHOR + NAD(+) =  
 CC 2,5-DIKETOCAMPHOR + NADH.  
 CC -1- COFACTOR: REQUIRES ZINC FOR ITS ACTIVITY.  
 CC -1- PATHWAY: SECOND STEP FOR CATABOLISM OF CAMPHOR.  
 CC -1- INDUCTION: BY CAMPHOR.  
 CC -1- SIMILARITY: BELONGS TO THE ZINC-CONTAINING ALCOHOL DEHYDROGENASE  
 CC FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: D14680; BAA03511.1; -  
 DR EMBL: M13471; AAA25762.1; ALT\_SEQ.  
 DR PIR: A29844; A29844.  
 DR HSP: P07846; LSDG.  
 DR PROSITE: PS00059; ADL\_ZINC; 1.  
 DR PFAM: PF00107; adh\_zinc; 1.  
 KW Oxidoreductase; Zinc; NAD.  
 FT METAL 40 40  
 FT METAL 62 62  
 FT METAL 98 98  
 FT METAL 101 101  
 FT METAL 104 104  
 FT METAL 170 170  
 FT METAL 170 170  
 SQ SEQUENCE 361 AA; 38460 MW; E46D28F7 CRC32;  
 ZINC (CATALYTIC) (BY SIMILARITY).  
 ZINC (CATALYTIC) (BY SIMILARITY).  
 ZINC (SECOND ATOM) (BY SIMILARITY).  
 ZINC (SECOND ATOM) (BY SIMILARITY).  
 ZINC (SECOND ATOM) (BY SIMILARITY).  
 ZINC (CATALYTIC) (BY SIMILARITY).  
 ZINC (CATALYTIC) (BY SIMILARITY).  
 Query Match 41.8%; Score 41; DB 1; Length 361;  
 Best Local Similarity 47.4%; Pred. No. 15;  
 Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;  
 QY 2 RLAIRLNRYRLAIRLNER 20  
 :|||:|||||:|:  
 DB 316 QLAARLDQRYPLADLTQR 334  
 RESULT 6  
 65KD\_ZYMO  
 ID 65KD\_ZYMO STANDARD; PRT; 584 AA.  
 AC P15255;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE 65 KD PROTEIN (ORF 1).  
 OS Zymomonas mobilis.  
 OG Plasmid pZM2.  
 OC Bacteria; Proteobacteria; alpha subdivision; Zymomonas group;  
 CC Zymomonas.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 10988 / ZM1;  
 RA MISAWA N., NAKAMURA K.;  
 RT "The nucleotide sequence of the 2.7 kilobase pair plasmid of Zymomonas



ID AC CYPI\_YEAST STANDARD; PRT; 1483 AA.  
 DT 01-OCT-1989 (Rel. 12, Created)  
 DT 01-OCT-1989 (Rel. 12, Last sequence update)  
 DT 15-DEC-1999 (Rel. 39, Last annotation update)  
 DE CYPI ACTIVATORY PROTEIN.  
 GN CYPI OR HAP1 OR YLR256W.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Hemiascomycetes; Saccharomycetales;  
 OC Saccharomycetaceae; Saccharomycetes.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 89125585.  
 RA CREUSOT F., VERDIERE J., GAISNE M., SLONIMSKI P.P.;  
 RT "CYPI (HAP1) regulator of oxygen-dependent gene expression in yeast.  
 RT I. Overall organization of the protein sequence displays several  
 RT novel structural domains.;  
 RL J. Mol. Biol. 204:263-276(1988).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 89106221.  
 RA PEIFFER K., KIM K.-S., KOGAN S., GUARENTE L.;  
 RT "Functional dissection and sequence of yeast HAP1 activator.";  
 RL Cell 56:291-301(1989).  
 RN [3]  
 RP MUTANT CYPI-18.  
 RX MEDLINE; 89125586.  
 RA VERDIERE J., GAISNE M., GUIARD B., DEFRAUX N., SLONIMSKI P.P.;  
 RT "CYPI (HAP1) regulator of oxygen-dependent gene expression in yeast.  
 RT II. Missense mutation suggests alternative Zn fingers as  
 RT discriminating agents of gene control.";  
 RL J. Mol. Biol. 204:277-282(1988).  
 RN [4]  
 RP DNA-BINDING.  
 RX MEDLINE; 97042358.  
 RA ZHANG L., GUARENTE L.;  
 RT "The C6 zinc cluster dictates asymmetric binding by HAP1.";  
 RL EMBO J. 15:4676-4681(1996).  
 RN [5]  
 RP STRUCTURE BY NMR OF 60-100.  
 RX MEDLINE; 96275662.  
 RA TIMMERMAN J., VUIDOT A.-L., BONTEMS F., LALLEMAND J.-Y., GERVAIS M.,  
 RA SHECHTER E., GUIARD B.;  
 RT "1H, 15N resonance assignment and three-dimensional structure of CYPI  
 RT (HAP1) DNA-binding domain.";  
 RL J. Mol. Biol. 259:792-804(1996).  
 CC -!- FUNCTION: REGULATION OF OXYGEN DEPENDENT GENE EXPRESSION. IT  
 CC MODULATES THE EXPRESSION OF ISO-1 AND ISO-2 CYTOCHROME C. BINDS  
 CC TO THE SEQUENCE 5'-CGGNNNTNCGG-3'.  
 CC -!- SUBUNIT: BINDS DNA AS A HOMODIMER.  
 CC -!- SUBCELLULAR LOCATION: NUCLEAR.  
 CC -!- MISCELLANEOUS: CYPI-18 MUTANT ACTIVATES THE EXPRESSION OF CYP3  
 CC (ISO-2) WHILE REDUCING THAT OF CYP1 (ISO-1).  
 CC -!- MISCELLANEOUS: HEME IS AN EFFECTOR MOLECULE FOR CYPI/HAP1. THE  
 CC REPEAT REGION (SEE FT TABLE) MEDIATES HEME INDUCTION BY MASKING  
 CC THE DNA-BINDING DOMAIN IN THE ABSENCE OF INDUCER.  
 CC -!- SIMILARITY: CONTAINS A ZN(2)-CYS(6), FUNGAL-TYPE BINUCLEAR  
 CC CLUSTER DOMAIN.  
 CC -!- SIMILARITY: CONTAINS 7 HEME REGULATORY MOTIFS (HRM).  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 DR EMBL; X13793; CAA32032.1; -.  
 DR EMBL; J03152; AAA34662.1; -.  
 DR PIR; S05804; RBYH1.  
 DR PDB; 1PYC; 01-AUG-96.  
 DR TRANSFAC; T00346; -.

SGD; L0002665; HAP1.  
 DR PROSITE; PS00463; ZN2\_Cy6\_FUNGAL\_1; 1.  
 DR PROSITE; PS00468; ZN2\_Cy6\_FUNGAL\_2; 1.  
 DR PFAM; PF00172; Zn\_clus; 1.  
 KW Transcription regulation; Activator; DNA-binding; Nuclear protein;  
 FT zinc; Metal-binding; Repeat; Heme; 3D-structure.  
 FT DNA\_BIND 64 93  
 FT ZN(2)-CYS(6), FUNGAL-TYPE.  
 FT VARIANT 63 63  
 FT S -> R (IN CYPI-18 MUTANT).  
 FT POLY-GLN.  
 FT DOMAIN 177 189  
 FT HRM 1 (POTENTIAL).  
 FT DOMAIN 280 285  
 FT HRM 2 (POTENTIAL).  
 FT DOMAIN 323 328  
 FT HRM 3 (POTENTIAL).  
 FT DOMAIN 347 352  
 FT HRM 4 (POTENTIAL).  
 FT DOMAIN 389 394  
 FT HRM 5 (POTENTIAL).  
 FT DOMAIN 415 420  
 FT HRM 6 (POTENTIAL).  
 FT DOMAIN 1192 1197  
 FT HRM 7 (POTENTIAL).  
 FT CONFLICT 145 145  
 FT T -> I (IN REF. 2).  
 FT CONFLICT 323 323  
 FT K -> R (IN REF. 2).  
 FT CONFLICT 455 455  
 FT S -> N (IN REF. 2).  
 FT CONFLICT 508 508  
 FT V -> M (IN REF. 2).  
 FT CONFLICT 587 587  
 FT N -> K (IN REF. 2).  
 FT CONFLICT 883 883  
 FT D -> N (IN REF. 2).  
 FT CONFLICT 960 960  
 FT H -> S (IN REF. 2).  
 FT CONFLICT 1151 1151  
 FT D -> N (IN REF. 2).  
 FT CONFLICT 1157 1157  
 FT S -> P (IN REF. 2).  
 FT CONFLICT 1305 1305  
 FT N -> Y (IN REF. 2).  
 FT CONFLICT 1483 1483  
 FT MW; 164150 MW; A46C7260 CRC32;  
 SQ SEQUENCE 1483 AA; 164150 MW; A46C7260 CRC32;  
 QY 2 RLAIRLNERYRLAIRLNE 19  
 DB 1124 KLTKQLSKKYNIAIRNKK 1141  
 Query Match 40.8%; Score 40; DB 1; Length 1483;  
 Best Local Similarity 38.9%; Pred. No. 1e+02;  
 Matches 7; Conservative 7; Mismatches 4; Indels 0; Gaps 0;  
 RESULT 10  
 YOR3\_BACCE  
 ID YOR3\_BACCE STANDARD; PRT; 312 AA.  
 AC O31352;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE HYPOTHETICAL PROTEIN (ORF3) (FRAGMENT).  
 OS Bacillus cereus.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
 OC Bacillus/Staphylococcus group; Bacillus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 10987;  
 RA KOLSTO A.B., OKSTAD O.A., LINDBACK T., HEGNA I., LAGREID A.,  
 RA RISHOVD A.L.;  
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
 CC -!- SIMILARITY: BELONGS TO THE ATZ/TRZ FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 DR EMBL; Y11139; CAA72025.1; -.  
 KW Hypothetical protein; Hydrolase.  
 FT NON\_TER 1  
 SQ SEQUENCE 312 AA; 35208 MW; 165E9409 CRC32;  
 Query Match 39.8%; Score 39; DB 1; Length 312;  
 Best Local Similarity 39.1%; Pred. No. 26;



DNA replication; Zinc-finger; ATP-binding; Cell cycle.  
 KW ZN\_FING 341 367 C4-TYPE (POTENTIAL).  
 FT DOMAIN 493 700 MCM.  
 FT NP\_BIND 543 550 ATP (POTENTIAL).  
 FT MUTAGEN 364 364 C->Y,F,S,H: LOSS OF ACTIVITY.  
 FT MUTAGEN 367 367 C->Y,F,S,H: LOSS OF ACTIVITY.  
 FT VARIANT 392 392 E -> K (IN STRAIN MCM2-1).  
 FT CONFLICT 164 164 S -> T (IN REF. 1).  
 FT CONFLICT 172 173 MD -> IH (IN REF. 1).  
 FT CONFLICT 529 529 G -> P (IN REF. 1).  
 FT CONFLICT 578 578 A -> R (IN REF. 1).  
 FT CONFLICT 583 583 D -> H (IN REF. 1).  
 FT CONFLICT 712 712 E -> O (IN REF. 1).  
 FT CONFLICT 733 747 MISSING (IN REF. 1).  
 FT CONFLICT 859 868 RSFAIYLGH -> SLQFIPWTKLLFLRISGYEDKKFS  
 VSIHVLAIFSIYKPFUFFV (IN REF. 1).  
 SQ SEQUENCE 868 AA; 98779 MW; C25586EA CRC32;  
 Query Match 39.8%; Score 39; DB 1; Length 868;  
 Best Local Similarity 50.0%; Pred. No. 82;  
 Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;  
 Qy 5 IRLNERYRLAIRLNER 20  
 Db 104 LSLERRIDAQLNER 119  
 RESULT 13  
 TEGU\_HSV6G STANDARD; PRT; 2077 AA.  
 AC P30002;  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE LARGE TEGUMENT PROTEIN.  
 GN U31.  
 OS Herpes simplex virus (type 6 / strain GS).  
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
 OC Betaherpesvirinae; Roseolovirus.  
 RN [1]  
 RX MEDLINE; 91374623.  
 RA JOSEPHS S.F., ABLASHI D.V., SALAHUDDIN S.Z., JAGODZINSKI L.L.,  
 RA WONG-STALF F., GALLO R.C.;  
 RT Identification of the human herpesvirus 6 glycoprotein H and  
 RT putative large tegument protein genes.";  
 RL J. Virol. 65:5597-5604(1991).  
 CC -!- FUNCTION: TEGUMENT PROTEIN.  
 CC -!- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL36,  
 CC EHV-1 24, EBV BPLF1, HVS-1 64, VZV 22, AND HCMV UL48.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; X83413; CAA58411.1;  
 DR EMBL; X83413; CAA58411.1;  
 SQ SEQUENCE 2077 AA; 239909 MW; 82ACA5DE CRC32;  
 Query Match 39.8%; Score 39; DB 1; Length 2077;  
 Best Local Similarity 50.0%; Pred. No. 2.2e+02;  
 Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;  
 Qy 9 ERYRLAIRLNER 20  
 Db 493 EKYKVALLNEK 504  
 RESULT 15  
 NTRB\_ECOLI STANDARD; PRT; 349 AA.  
 AC P06712;  
 DT 01-JAN-1988 (Rel. 06, Created)  
 DT 01-JAN-1988 (Rel. 06, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE NITROGEN REGULATION PROTEIN NR(II) (EC 2.7.3.-).  
 GN GNLR OR NTRB OR GLNR.  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-K12;  
 RX MEDLINE; 87174797.  
 RA MIRANDA-RIOS J., SANCHEZ-PESCADOR R., URDEA M., COVARRUBIAS A.A.;  
 RT "The complete nucleotide sequence of the glnALG operon of Escherichia  
 RT coli K12.";  
 RL Nucleic Acids Res. 15:2757-2770(1987).  
 CC [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-K12 / MG1655;  
 RX MEDLINE; 93347969.  
 RA PLUNKETT G. III,  
 RT "Analysis of the Escherichia coli genome. III. DNA sequence of the  
 RT region from 87.2 to 89.2 minutes.";

RESULT 14  
 TEGU\_HSV6U STANDARD; PRT; 2077 AA.  
 AC P52340;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE LARGE TEGUMENT PROTEIN.  
 GN U31 OR HHRF1.  
 OS Herpes simplex virus (type 6 / strain Uganda-1102).  
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
 OC Betaherpesvirinae; Roseolovirus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 95266321.  
 RA GOMPELS U.A., NICHOLAS J., LAWRENCE G., JONES M., THOMSON B.J.,  
 RA MARTIN M.E., EFSTATHIOU S., CRAXTON M., MACAULAY H.A.;  
 RT "The DNA sequence of human herpesvirus-6: structure, coding content,  
 RT and genome evolution.";  
 RL Virology 209:29-51(1995).  
 CC -!- FUNCTION: TEGUMENT PROTEIN.  
 CC -!- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL36,  
 CC EHV-1 24, EBV BPLF1, HVS-1 64, VZV 22, AND HCMV UL48.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; X83413; CAA58411.1;  
 DR EMBL; X83413; CAA58411.1;  
 SQ SEQUENCE 2077 AA; 239946 MW; 36FCF7B1 CRC32;  
 Query Match 39.8%; Score 39; DB 1; Length 2077;  
 Best Local Similarity 50.0%; Pred. No. 2.2e+02;  
 Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;  
 Qy 9 ERYRLAIRLNER 20  
 Db 493 EKYKVALLNEK 504  
 RESULT 15  
 NTRB\_ECOLI STANDARD; PRT; 349 AA.  
 AC P06712;  
 DT 01-JAN-1988 (Rel. 06, Created)  
 DT 01-JAN-1988 (Rel. 06, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE NITROGEN REGULATION PROTEIN NR(II) (EC 2.7.3.-).  
 GN GNLR OR NTRB OR GLNR.  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-K12;  
 RX MEDLINE; 87174797.  
 RA MIRANDA-RIOS J., SANCHEZ-PESCADOR R., URDEA M., COVARRUBIAS A.A.;  
 RT "The complete nucleotide sequence of the glnALG operon of Escherichia  
 RT coli K12.";  
 RL Nucleic Acids Res. 15:2757-2770(1987).  
 CC [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-K12 / MG1655;  
 RX MEDLINE; 93347969.  
 RA PLUNKETT G. III,  
 RT "Analysis of the Escherichia coli genome. III. DNA sequence of the  
 RT region from 87.2 to 89.2 minutes.";

```

RL Nucleic Acids Res. 21:3391-3398(1993).
RN [3]
RP SEQUENCE OF 1-22 FROM N.A.
RX MEDLINE; 85006814.
RA UENO-NISHIO S., MANGO S., REITZER L.J., MAGASANIK B.;
RT "Identification and regulation of the glnL operator-promoter of the
RT complex glnALG operon of Escherichia coli.";
RL J. Bacteriol. 160:379-384(1984).
RN [4]
RP SEQUENCE OF 1-24 FROM N.A.
RX MEDLINE; 86031370.
RA ROCHA M., VAZQUEZ M., GARCIAARRUBIO A., COVARRUBIAS A.A.;
RT "Nucleotide sequence of the glnA-glnL intercistronic region of
RT Escherichia coli.";
RL Gene 37:91-99(1985).
RN [5]
RP PHOSPHORYLATION SITE.
RX MEDLINE; 91201336.
RA NINFA A.J., BENNETT R.L.;
RT "Identification of the site of autophosphorylation of the bacterial
RT protein kinase/phosphatase NRII.";
RL J. Biol. Chem. 266:6888-6893(1991).
CC -!- FUNCTION: NTRB ACTS AS A SIGNAL TRANSDUCER WHICH RESPONDS TO THE
CC NITROGEN LEVEL OF CELL AND MODULATES THE ACTIVITY OF NTRC. IN
CC NITROGEN LIMITATION NTRB ACTIVATES NTRC BY PHOSPHORYLATING IT,
CC WHILE IN NITROGEN EXCESS NTRC IS DEPHOSPHORYLATED AND CONSEQUENTLY
CC INACTIVATED BY NTRB.
CC -!- SIMILARITY: TO OTHER PROKARYOTIC SENSORY TRANSDUCTION HISTIDINE
CC KINASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X05173; CAA28807.1; -
DR EMBL; K02176; AAA23881.1; -
DR EMBL; L19201; AAB03003.1; -
DR EMBL; AE000462; AAC76866.1; -
DR PIR; Q00553; RGECL.
DR PIR; B23970; B23970.
DR PIR; S40814; S40814.
DR ECGENE; EGI0387; GLNL.
DR PFAM; PF00512; signal; 1.
KW Sensory transduction; Transferase; Kinase; Phosphorylation;
KW Nitrogen fixation; ATP-binding.
FT DOMAIN 116 349 TRANSMITTER DOMAIN (POTENTIAL).
FT MOD_RES 139 139 PHOSPHORYLATION (AUTO-).
FT BINDING 329 329 ATP (By SIMILARITY).
SQ SEQUENCE 349 AA; 38556 MW; 6A017919 CRC32;

```

```

Query Match          39.3%; Score 38.5; DB 1; Length 349;
Best Local Similarity 50.0%; Pred. No. 36;
Matches 10; Conservative 3; Mismatches 4; Indels 3; Gaps 1;

QY 2 RLAIRL---NERYFLAIRLN 18
   | | | | | | | | | |
Db 265 RTAQLTLHGERYFLAARID 284

```

Search completed: February 8, 2000, 01:26:01  
Job time: 1561 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2000, 19:16:14 ; Search time 176.54 Seconds  
(without alignments)  
7.855 Million cell updates/sec

Title: US-08-653-294-31

Perfect score: 98

Sequence: 1 YRLAIRLNERYRLAIRLNER 20

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 225878 seqs, 69334122 residues

Total number of hits satisfying chosen parameters: 225878

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database :

SPTREMBL\_12.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phage.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	44.5	45.4	191	12 Q98543	Q98543 parametium
2	44	44.9	219	5 Q19332	Q19332 caenorhabdi
3	44	44.9	2391	10 Q9X840	Q9X840 cryza sativ
4	42.5	43.4	414	5 Q19088	Q19088 caenorhabdi
5	42	42.9	188	8 Q35521	Q35521 phytophthor
6	42	42.9	623	5 Q9XAJ3	Q9XAJ3 streptomyc
7	41.5	42.3	2632	5 P90736	P90736 caenorhabdi
8	41	41.8	223	2 P71724	P71724 mycobacteri
9	41	41.8	413	2 O50931	O50931 borrelia bu
10	41	41.8	471	5 Q9Y1X6	Q9Y1X6 ephydatia f
11	40	40.8	192	2 O67035	O67035 aquifex aeo
12	40	40.8	319	5 Q94499	Q94499 dictyosteli
13	40	40.8	369	13 Q9W678	Q9W678 barbus barb
14	40	40.8	433	2 P71096	P71096 bacteroides
15	40	40.8	526	2 P74182	P74182 synecocyst
16	40	40.8	579	2 Q9WZ01	Q9WZ01 thermotoga
17	40	40.8	741	2 Q9XAU3	Q9XAU3 pseudomonas
18	40	40.8	1001	4 O75150	O75150 homo sapien
19	40	40.8	1164	5 P92021	P92021 caenorhabdi
20	40	40.8	1502	3 Q06574	Q06574 saccharomyc

21	39.5	40.3	540	2	P74663	P74663 synecocyst
22	39.5	40.3	1110	2	P95078	P95078 mycobacteri
23	39	39.8	143	8	Q9XMS8	Q9XMS8 tetrahymena
24	39	39.8	144	5	Q27369	Q27369 trypanosoma
25	39	39.8	268	2	O07709	O07709 mycobacteri
26	39	39.8	281	6	Q29475	Q29475 canis famil
27	39	39.8	320	2	O51569	O51569 borrelia bu
28	39	39.8	332	2	O69183	O69183 alcaligenes
29	39	39.8	871	12	O91276	O91276 chimpanzee
30	39	39.8	871	12	O91277	O91277 chimpanzee
31	39	39.8	888	4	O94946	O94946 homo sapien
32	39	39.8	1044	3	Q00943	Q00943 pichia angu
33	39	39.8	2078	12	O69055	O69055 human herpe
34	38.5	39.3	193	1	O9YEP7	O9YEP7 aeropyrum p
35	38	38.8	79	2	O50974	O50974 neisseria g
36	38	38.8	120	5	O97363	O97363 penaeus mon
37	38	38.8	172	2	O85216	O85216 actinobacil
38	38	38.8	226	3	O42852	O42852 schizosacch
39	38	38.8	229	5	Q19415	Q19415 caenorhabdi
40	38	38.8	270	11	O35655	O35655 mus musculu
41	38	38.8	289	2	O9WZ64	O9WZ64 thermotoga
42	38	38.8	297	2	O9Z3J1	O9Z3J1 rhizobium s
43	38	38.8	340	2	O48759	O48759 listeria mo
44	38	38.8	346	2	O48229	O48229 haemophilus
45	38	38.8	369	2	P95702	P95702 staphylococ

## ALIGNMENTS

RESULT 1

Q98543 ID Q98543 PRELIMINARY; PRT; 191 AA.  
AC Q98543; 1997 (TREMBLrel. 02, Created)  
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)  
DE 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
DE GENOME, PARTIAL SEQUENCE.  
GN A493L  
OS Parametium bursaria chlorella virus 1 (PBCV-1).  
OC Viruses; dsDNA viruses, no RNA stage; Phycodnaviridae; Phycodnavirus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE: 96400190.  
RA KUTISH G.F., LI Y., LU Z., FURUTA M., ROCK D.L., VAN ETTE J.L.;  
RT "Analysis of 76 kb of the chlorella virus PBCV-1 330-kb genome: map  
positions 182 to 258.";  
RL Virology 223:303-317(1996).  
DR EMBL; U42580; AAC96860.1; -.  
SQ SEQUENCE 191 AA; 22651 MW; E4547C83 CRC32;

Query Match 45.4%; Score 44.5; DB 12; Length 191;

Best Local Similarity 45.5%; Pred. No. 10;

Matches 10; Conservative 4; Mismatches 5; Indels 3; Gaps 1;

Qy 1 YRLAIRLNE---RYRLAIRLNE 19

Db 61 YKLEIRYKNGKRYLVVRENE 82

RESULT 2

Q19332 ID Q19332 PRELIMINARY; PRT; 219 AA.  
AC Q19332;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DE 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
DT 01-JAN-1999 (TREMBLrel. 09, Last annotation update)  
DE FILA1.2 PROTEIN.  
GN FILA1.2  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.

```

RN  SEQUENCE FROM N.A.
RP  HARRIS B.; (AUG-1995) to the EMBL/GenBank/DBJ databases.
RL  Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
RN  [2]
RX  MEDLINE: 94150718.
RA  WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,
RA  BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,
RA  CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
RA  GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA  JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
RA  LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA  PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA  SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,
RA  THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
RA  WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;
RT  "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT  elegans.";
RL  Nature 368:32-38(1994).
DR  EMBL: 250857; CAA90718.1; --
SQ  SEQUENCE 219 AA; 25040 MW; 156231A7 CRC32;

Query Match 44.9%; Score 44; DB 5; Length 219;
Best Local Similarity 61.5%; Pred. No. 14;
Matches 8; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY  5 IRLNERYRLAIRL 17
DB  133 VRYNKKRLAURL 145
:|:|:|:|:|:|

RESULT 3
QY  Q9XE40 PRELIMINARY; PRT; 2391 AA.
AC  Q9XE40;
DT  01-NOV-1999 (TREMREL. 12, Created)
DT  01-NOV-1999 (TREMREL. 12, Last sequence update)
DT  01-NOV-1999 (TREMREL. 12, Last annotation update)
DE  HYPOTHETICAL 27.9 KD PROTEIN.
OS  Oryza sativa (Rice).
OC  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC  euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales;
OC  Poaceae; Oryza.
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN-CV. NIPPONBARE;
RA  SASAKI T., NAGAMURA Y., YAMAMOTO K.;
RT  "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 6, PAC
RT  clone:P0680A03.";
RL  Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR  EMBL: A8023482; BAA78744.1; --
KW  HYPOTHETICAL PROTEIN.
SQ  SEQUENCE 2391 AA; 278546 MW; EBD126F9 CRC32;

Query Match 44.9%; Score 44; DB 10; Length 2391;
Best Local Similarity 57.1%; Pred. No. 1.8e+02;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY  6 RLNERYRLAIRLNE 19
DB  870 RLKESYSAVRLNQ 883
:|:|:|:|:|:|

RESULT 4
QY  Q19088 PRELIMINARY; PRT; 414 AA.
AC  Q19088;
DT  01-NOV-1996 (TREMREL. 01, Created)
DT  01-NOV-1996 (TREMREL. 01, Last sequence update)
DT  01-MAY-1999 (TREMREL. 10, Last annotation update)

```

```

DE  SIMILARITY TO RODLIKE TAIL DOMAIN OF MYOSIN HEAVY CHAIN.
CN  F01F1.4.
OS  Caenorhabditis elegans.
OC  Eukaryota; Metazoa; Nematoda; Secernentea; Rhabdita; Rhabditida;
OC  Rhabditina; Rhabditoides; Rhabditidae; Peloderinae; Caenorhabditis.
RN  [1]
RX  MEDLINE: 94150718.
RA  WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,
RA  BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,
RA  CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
RA  GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA  JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
RA  LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA  PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA  SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,
RA  THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
RA  WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;
RT  "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT  elegans.";
RL  Nature 368:32-38(1994).
RN  [2]
RP  SEQUENCE FROM N.A.
RC  STRAIN-BRISTOL N2;
RA  WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,
RA  BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,
RA  CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
RA  GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA  JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
RA  LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA  PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA  SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,
RA  THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
RA  WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;
RT  "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT  elegans.";
RL  Nature 368:32-38(1994).
RN  [3]
RP  SEQUENCE FROM N.A.
RC  STRAIN-BRISTOL N2;
RA  MILLER N.;
RL  Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.
RN  [3]
RP  SEQUENCE FROM N.A.
RC  STRAIN-BRISTOL N2;
RA  WATERSTON R.;
RL  Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
DR  EMBL: U13070; AAC46638.1; --
DR  PFW; PFO1363; FIVE; 1.
SQ  SEQUENCE 414 AA; 47599 MW; 53B60FB7 CRC32;

Query Match 43.4%; Score 42.5; DB 5; Length 414;
Best Local Similarity 52.6%; Pred. No. 48;
Matches 10; Conservative 4; Mismatches 4; Indels 1; Gaps 1;

QY  2 RLAIRLNERYRLAIRLNER 20
DB  62 KAAMELNERYLQDLS-NER 79
:|:|:|:|:|:|

RESULT 5
QY  Q35521 PRELIMINARY; PRT; 188 AA.
AC  Q35521;
DT  01-NOV-1996 (TREMREL. 01, Created)
DT  01-NOV-1996 (TREMREL. 01, Last sequence update)
DT  01-NOV-1999 (TREMREL. 12, Last annotation update)
DE  NADH UBIQUINONE OXIDOREDUCTASE SUBUNIT.
OS  Phytophthora megasperma (Potato pink rot fungus).
OC  Mitochondrion.
OC  Eukaryota; stramenopiles; Oomycetes; Peronosporales; Pythiaceae;
OC  Phytophthora.
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN-695T; TISSUE-MYCELIUM;
RA  CLARY S.A., HUDSPETH M.E.;
RL  Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.
RN  [2]
RP  SEQUENCE FROM N.A.
RC  STRAIN-695T; TISSUE-MYCELIUM;
RA  SACHAY D.J., HUDSPETH D.S.S., NADLER S.A., HUDSPETH M.E.S.;
RL  Exp. Mycol. 17:7-23(1993).
DR  EMBL: L16863; CAB24134.1; --
DR  PROSITE: PS00542; COMPLEX1_30K; 1.
DR  PFW; PFO0329; complex1_30Kd; 1.
KW  Ubiquinone; Mitochondrion.
SQ  SEQUENCE 188 AA; 22638 MW; 61C1339B CRC32;

```

Submitted (OCT-1996) to the EMBL/GenBank/DDBJ databases.  
[2]  
RN REVISIONS.  
RP STRAIN-BRISTOL N2;  
RA JONES S.J.M.;  
RL Submitted (MAY-1997) to the EMBL/GenBank/DDBJ databases.  
[3]  
RN SEQUENCE FROM N.A.  
RA PERCY C.;  
RL Submitted (OCT-1996) to the EMBL/GenBank/DDBJ databases.  
CC -!- SIMILARITY: WPAK TO DYNEIN HEAVY CHAINS.  
DR EMBL; Z81028; CAB02695.1; -.  
DR EMBL; Z81096; CAB02695.1; JOINED.  
DR EMBL; Z81096; CAB03163.1; -.  
DR EMBL; Z81028; CAB03163.1; JOINED.  
DR WORPEP; B0365.7; CE07724.  
KW Hypothetical protein.  
SQ SEQUENCE 2632 AA; 305774 MW; 4F3356EF CRC32;

Query Match 42.3%; Score 41.5; DB 5; Length 2632;  
Best Local Similarity 43.5%; Pred. No. 4.9e+02;  
Matches 10; Conservative 5; Mismatches 5; Indels 3; Gaps 1;

QY 1 YRLAIRLNERYRLAIR---LNER 20  
|||::|| ::| :| ||  
DB 557 YRSVRVINEVKVQIKMGNLEER 579

RESULT 8  
P71724 PRELIMINARY; PRT; 223 AA.

ID P71724 AC F71724;  
DT 01-FEB-1997 (TrEMBLrel. 02, Created)  
DI 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)  
DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)  
DE HYPOTHETICAL 24.2 KD PROTEIN.  
GN RV2419C.  
OS Mycobacterium tuberculosis.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
CN [1]  
SQ SEQUENCE FROM N.A.  
STRAIN=H37RV;  
RX MEDLINE; 98295987.  
RA COLE S.T., BROSCH R., PARKHILL J., GARNIER T., CHURCHER C., HARRIS D.,  
RA GORDON S.V., EIGLMEIER K., GAS S., BARRY III C.E., TEKAIA F.,  
RA BADCOCK K., BASHAM D., BROWN D., FELTWELL T., CONNOR R.,  
RA DAVIES R., DEVLIN K., FELTWELL T., GENTLES S., HAMLIN N., HOLROYD S.,  
RA HORNSBY T., JAGELS K., KROGH A., MCLEAN J., MOULE S., MURPHY L.,  
RA OLIVER S., OSBORNE J., QUAIL M.A., RAJANDREAM M.A., ROGERS J.,  
RA RUTHER S., SEEGER K., SKELTON S., SOARES S., SOARES R., SULSTON J.E.,  
RA TAYLOR K., WHITEHEAD S., BARRELL B.G.;  
RA "Deciphering the biology of Mycobacterium tuberculosis from the  
RT complete genome sequence.";  
RL Nature 393:537-544(1998).  
RN [2]  
SQ SEQUENCE FROM N.A.  
STRAIN=H37RV;  
RX PARKHILL J.;  
RA Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases.  
RL EMBL; Z81368; CAB03751.1; -.  
DR HSSP; P00950; 3PGM.  
DR PFAM; PF00300; PGAM; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 223 AA; 24174 MW; CECA2812 CRC32;

Query Match 41.8%; Score 41; DB 2; Length 223;  
Best Local Similarity 40.0%; Pred. No. 42;  
Matches 8; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 YRLAIRLNERYRLAIRLNER 20

Db 62 YDTAVKLGRTGLVRVDTR 81  
 | | : | | | : | : | : |

RESULT 9  
 ID O50931 PRELIMINARY; PRT; 413 AA.  
 AC O50931;  
 DT 01-JUN-1998 (TREMBlrel. 06, Created)  
 DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)  
 DE HYPOTHETICAL 47.9 KD PROTEIN.  
 GN BBA38.  
 OS Borrelia burgdorferi (Lyme disease spirochete).  
 OG Plasmid lp54  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 RN [1]

SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 35210 / B31;  
 RX MEDLINE: 98065943.  
 RA FRASER C.M., CASJENS S., HUANG W.M., SUTTON G.G., CLAYTON R.A.,  
 RA LAHIGRA R., WHITE O., KETCHUM K.A., DODSON R., HICKEY E.K., GWINN M.,  
 RA DOUGHERTY B., TOMB J.-F., FLEISCHMANN R.D., RICHARDSON D.,  
 RA PETERSON J., KERLAVAGE A.R., QUACKENBUSH J., SALZBERG S., HANSON M.,  
 RA VAN VUGT R., PALMER N., ADAMS M.D., GOCAYNE J.D., WEIDMAN J.,  
 RA UTTERBACK T., WATTHEY L., MCDONALD L., ARTIACH P., BOWMAN C.,  
 RA GARLAND S., FUJII C., COTTON M.D., HORST K., ROBERTS K., HATCH B.,  
 RA SMITH H.O., VENTER J.C.;  
 RT "Genomic sequence of a Lyme disease spirochaete, Borrelia  
 burgdorferi."  
 RL Nature 390:580-586(1997).  
 DR EMBL: AE000790; AAC66300.1; -;  
 DR TIGR: BBA38; -;  
 KW Hypothetical protein; Plasmid.  
 SQ SEQUENCE 413 AA; 47901 MW; 5959431C CRC32;

Query Match 41.8%; Score 41; DB 2; Length 413;  
 Best Local Similarity 46.7%; Pred. No. 81;  
 Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Oy 5 IRLNERYRLAIRLNE 19  
 : | : | : | : | : | : |

Db 351 LKLNKYRLNKNFNE 365

RESULT 10  
 ID Q9Y1X6 PRELIMINARY; PRT; 471 AA.  
 AC Q9Y1X6;  
 DT 01-NOV-1999 (TREMBlrel. 12, Created)  
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
 DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)  
 DE SPTP4 (EC 3.1.3.48) (FRAGMENT).  
 OS Ephydatia fluviatilis.  
 OC Eukaryota; Metazoa; Porifera; Demospongiae; Ceractinomorpha;  
 OC Haplosclerida; Spongillidae; Ephydatia.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 99246376.  
 RA ONO K., IWABE N., KUMA K., MIYATA T.;  
 RT "Multiple protein tyrosine phosphatases in sponges and explosive gene  
 duplication in the early evolution of animals before the parazoan-  
 eumetazoan split."  
 RL J. Mol. Evol. 48:654-662(1999).  
 DR EMBL: AB019125; BAA82558.1; -;  
 DR PROSITE; PS00383; TYR\_PHOSPHATASE\_1; 2.  
 KW Hydrolase.  
 FT NON\_TER 1  
 SQ SEQUENCE 471 AA; 54509 MW; BA496490 CRC32;

Query Match 41.8%; Score 41; DB 5; Length 471;

Best Local Similarity 61.5%; Pred. No. 94;  
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Oy 7 IRLNERYRLAIRLNE 19  
 | | : | | | : | : | : |

Db 181 LNEQYRLCALND 193

RESULT 11  
 ID O67035 PRELIMINARY; PRT; 192 AA.  
 AC O67035;  
 DT 01-AUG-1998 (TREMBlrel. 07, Created)  
 DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)  
 DE HYPOTHETICAL 22.5 KD PROTEIN.  
 GN AQ\_880.  
 OS Aquifex aeolicus.  
 OC Bacteria; Aquificales; Aquificaceae; Aquifex.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-VF5;  
 RX MEDLINE: 98196666.  
 RA DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,  
 RA GRAHAM D.E., OVERBECK R., SNEAD M.A., KELLER M., AUJAY M., HUBER R.,  
 RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;  
 RT "The complete genome of the hyperthermophilic bacterium Aquifex  
 aeolicus."  
 RL Nature 392:353-358(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-VF5;  
 RA DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,  
 RA GRAHAM D.E., OVERBECK R., SNEAD M.A., KELLER M., AUJAY M., HUBER R.,  
 RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;  
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AE000712; AAC07003.1; -;  
 KW Hypothetical protein.  
 SQ SEQUENCE 192 AA; 22532 MW; BD174502 CRC32;

Query Match 40.8%; Score 40; DB 2; Length 192;  
 Best Local Similarity 53.3%; Pred. No. 51;  
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Oy 2 RLAILNERYRLAIR 16  
 : | : | : | : | : | : |

Db 119 KLIALRLKRYREALR 133

RESULT 12  
 ID Q94499 PRELIMINARY; PRT; 319 AA.  
 AC Q94499;  
 DT 01-FEB-1997 (TREMBlrel. 02, Created)  
 DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)  
 DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)  
 DE KINESIN-LIKE PROTEIN K6 (FRAGMENT).  
 OS Dictyostelium discoideum (Slime mold).  
 OC Eukaryota; Dictyostelida; Dictyostelium.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-AX3;  
 RA DEHOSTOS E.L., MCCAFFREY G., VALE R.D.;  
 RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: U69984; AAB09082.1; -;  
 DR HSSP; P17119; 3KAR.  
 DR PFAM; PF00225; kinesin; 1.  
 FT NON\_TER 1  
 SQ SEQUENCE 319 AA; 35971 MW; D96F9939 CRC32;

Query Match 40.8%; Score 40; DB 5; Length 319;  
 Best Local Similarity 69.2%; Pred. No. 88;  
 Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 7 LNERYLRLAIRLNE 19  
 |||| :|| :||  
 DB 160 LNERKKLACRENE 172

## RESULT 13

ID Q9W678 PRELIMINARY; PRT; 369 AA.  
 AC Q9W678;  
 DT 01-NOV-1999 (TREMELrel. 12, Created)  
 DT 01-NOV-1999 (TREMELrel. 12, Last sequence update)  
 DT 01-NOV-1999 (TREMELrel. 12, Last annotation update)  
 DE CELLULAR TUMOR ANTIGEN P53.  
 GN P53.  
 OS Barbus barbus (barbel).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;  
 OC Neoterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinoidae; Cyprinidae; Cyprininae; Barbus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA BHASKARAN A., MAY D., RAND-WEAVER M., TYLER C.R.;  
 RT "Evolutionary conservancy of p53 gene sequences in fish."  
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: P53 SEEMS TO ACT AS A TUMOR SUPPRESSOR IN SOME, BUT  
 CC PROBABLY NOT ALL, TUMOR TYPES. P53 IS PROBABLY INVOLVED IN CELL  
 CC CYCLE REGULATION, IT IS A TRANS-ACTIVATOR THAT ACTS TO NEGATIVELY  
 CC REGULATE CELLULAR DIVISION BY CONTROLLING A SET OF GENES REQUIRED  
 CC FOR THIS PROCESS. ONE OF THE GENES ACTIVATED IS AN INHIBITOR OF  
 CC CYCLIN-DEPENDENT KINASES (BY SIMILARITY).  
 CC -!- SUBCELLULAR LOCATION: NUCLEAR.  
 DR EMBL; AF071570; AAD34212.1; -.  
 DR PROSITE; PS00348; P53; 1.  
 KW Anti-oncogene; DNA-binding; Transcription regulation; Activator;  
 KW Nuclear protein; Phosphorylation.  
 SQ SEQUENCE 369 AA; 41233 MW; 0D4E77F7 CRC32;

Query Match 40.8%; Score 40; DB 13; Length 369;  
 Best Local Similarity 36.8%; Pred. No. 1e+02;  
 Matches 7; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 YRLAIRLNERYLRLAIRLNE 19  
 || :|| :|| :||  
 DB 301 YTLQVRGKERYEMKKIND 319

## RESULT 14

ID P71096 PRELIMINARY; PRT; 433 AA.  
 AC P71096;  
 DT 01-FEB-1997 (TREMELrel. 02, Created)  
 DT 01-FEB-1997 (TREMELrel. 02, Last sequence update)  
 DT 01-AUG-1998 (TREMELrel. 07, Last annotation update)  
 DE REGULATORY PROTEIN.  
 GN SUSA.  
 OS Bacteroides thetaiotaomicron.  
 OC Bacteria; Cytophagales; Bacteroidaceae; Bacteroides.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-5842;  
 RA D'ELIA J., SALYERS A.;  
 RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U67062; AAB39215.1; -.  
 SQ SEQUENCE 433 AA; 50646 MW; 13D3D4D5 CRC32;

Query Match 40.8%; Score 40; DB 2; Length 433;  
 Best Local Similarity 52.6%; Pred. No. 1.2e+02;  
 Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 YRLAIRLNERYLRLAIRLNE 19  
 || :|| :|| :|| :|| :|| :||  
 DB 407 YSLATIYNSRIAIKLPE 425

## RESULT 15

P74182 PRELIMINARY; PRT; 526 AA.  
 ID P74182;  
 AC P74182;  
 DT 01-FEB-1997 (TREMELrel. 02, Created)  
 DT 01-FEB-1997 (TREMELrel. 02, Last sequence update)  
 DT 01-JAN-1999 (TREMELrel. 09, Last annotation update)  
 DE HYPOTHETICAL 57.7 KD PROTEIN.  
 OS Synechocystis sp. (strain PCC 6803).  
 OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-PCC6803;  
 RA TABATA S.;  
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-PCC6803;  
 RX MEDLINE; 97061201.  
 RA KANEKO T., SATO S., KOTANI H., TANAKA A., ASAMIZU E., NAKAMURA Y.,  
 RA MIYAJIMA T., HIROSAWA M., SUGIURA M., SASAMOTO S., KIMURA T.,  
 RA HOSOUCHI N., MURAKI A., NAKAZAKI N., NARUO K., OKUMURA S.,  
 RA SHIMPO S., TAKEUCHI C., WADA T., WATANABE A., YAMADA M., YASUDA M.,  
 RA TABATA S.;  
 RT "Sequence analysis of the genome of the unicellular cyanobacterium  
 RT Synechocystis sp. strain PCC6803. II. Sequence determination of the  
 RT entire genome and assignment of potential protein-coding regions."  
 RL DNA Res. 3:109-136(1996).  
 DR EMBL; D90913; BAA18271.1; -.  
 KW Hypothetical protein.  
 SQ SEQUENCE 526 AA; 57735 MW; 7D6F9B34 CRC32;

Query Match 40.8%; Score 40; DB 2; Length 526;  
 Best Local Similarity 52.6%; Pred. No. 1.5e+02;  
 Matches 10; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 RLAIRLNERYLRLAIRLNER 20  
 || :|| :|| :|| :|| :|| :||  
 DB 98 ROAITLDEAIDLALRNNEQ 116

Search completed: February 8, 2000, 19:16:17  
 Job time: 21504 sec

**THIS PAGE BLANK (USPTO)**

OM of: US-08-653-294-31 to: GenEmbl.\* out\_format : pfs  
 Date: Feb 8, 2000 10:23 PM  
 About: Results were produced by the GenCore software, version 4.5,  
 Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:  
 -MODEL=framet\_p2n.model -DEV=xlpr  
 -Q=/cnrl\_1/USPro\_spool/US08653294/runat\_04022000\_160701\_15779/app\_query.fasta.2  
 -DB=GenEmbl -QEMT=fastap -SUFFIX=rge -GAPOP=12.000 -GAPEXT=4.500  
 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -QGAPOP=4.500  
 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000  
 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000  
 -DELEXT=7.000 -START=1 -MATRIX=blossum62 -TRANS=human40.cdi  
 -LIST=45 -LOCALIGN=200 -THR\_SCORE=pct -ALIGN=15 -MODE=LOCAL  
 -OUTMT=pfs -NORM=ext -MINLEN=0 -MAXLEN=1000000 -USER=US08653294  
 -NCPU=6 -ICPU=3 -NO\_XLPXY -WAIT -THREADS=1

## Search information block:

Query: US-08-653-294-31

Query length: 20

Database: GenEmbl.\*

Database sequences: 821193

Database length: 1518192014

Search time (sec): 10176.920000

## score\_list:

Sequence	Strd Orig	Zscore	Escore	Len	Documentation
gb_hn1:CEY38E10_3	-	55.00	113.77	258.55	93707 ! Continuation (4 of 4) of CEY3
gb_hn1:CEY38E10A	-	55.00	112.23	314.83	110932 ! AL110484 Caenorhabditis eleg
gb_hn1:CEY41C4A	+	54.00	107.73	560.95	129060 ! AL032627 Caenorhabditis eleg
gb_hn1:CEY68A7	-	54.00	103.85	922.97	197735 ! AL022282 Caenorhabditis eleg
gb_hn1:CEY41C4	+	54.00	101.32	1.3e+03	261002 ! AL022278 Caenorhabditis eleg
gb_hn1:CEY51A2_3	-	53.00	106.74	636.81	130249 ! Continuation (4 of 4) of CEY
gb_hn1:CEY51A2D	-	53.00	103.91	915.37	139259 ! AL021497 Caenorhabditis eleg
gb_hn1:CEY06889	-	53.00	97.99	2.0e+03	267118 ! AC006889 Caenorhabditis eleg
gb_hn1:AC0068792	-	53.00	96.95	2.2e+03	299202 ! AC006792 Caenorhabditis eleg
gb_hn1:AC003070	-	52.00	101.42	1.3e+03	129914 ! AC003070 Homo sapiens chrom
gb_hn1:CEY4764	-	51.00	112.55	302.25	27111 ! AF043694 Caenorhabditis eleg
gb_hn1:CEY52B11A	+	51.00	110.58	389.26	33673 ! AC006672 Caenorhabditis eleg
gb_hn1:CEY37D8_0	+	51.00	90.71	5.0e+03	299015 ! AC006842 Caenorhabditis eleg
gb_hn1:CEY57A10_2	+	51.00	90.49	5.1e+03	306131 ! AC006874 Caenorhabditis eleg
gb_hn1:CEY44C11	-	50.00	117.07	169.41	11710 ! AF045637 Caenorhabditis eleg
gb_hn1:CEY4764	-	50.00	107.26	595.85	34397 ! Z99171 Caenorhabditis eleg
gb_hn1:CEY52B11A	+	50.00	96.94	2.2e+03	106949 ! AL032654 Caenorhabditis eleg
gb_hn1:CEY111B2_3	+	50.00	96.68	2.3e+03	110000 ! Continuation (4 of 7) of CEY
gb_hn1:CEY37D8_0	+	50.00	96.68	2.3e+03	110000 ! Z92819 Caenorhabditis eleg
gb_hn1:CEY57A10_2	+	50.00	96.68	2.3e+03	110000 ! Continuation (3 of 5) of CEY
gb_hn1:CEY57A10_2	+	50.00	96.68	2.3e+03	110000 ! Z98873 Caenorhabditis eleg
gb_hn1:CEY71A12_2	+	50.00	96.68	2.3e+03	110000 ! Continuation (3 of 4) of CEY
gb_hn1:CEY37D8A	+	50.00	94.97	2.9e+03	132742 ! AL032626 Caenorhabditis eleg
gb_hn1:CEY52B11	+	50.00	89.81	5.6e+03	233941 ! Z98867 Caenorhabditis eleg
gb_hn1:CEY068904	+	50.00	89.03	6.2e+03	254877 ! AC006904 Caenorhabditis eleg
gb_hn1:CEY43B3B	+	49.00	129.86	32.83	2036 ! AL032679 Caenorhabditis eleg
gb_hn1:CEY43B3B	+	49.00	105.22	773.88	30522 ! U70853 Caenorhabditis eleg
gb_hn1:CEY43B3B	+	49.00	103.75	934.64	35879 ! Z82286 Caenorhabditis eleg
gb_hn1:CEY02821	-	49.00	103.31	989.19	37666 ! Z81556 Caenorhabditis eleg
gb_hn1:CEY13640	+	49.00	100.04	1.5e+03	53960 ! AC013640 Homo sapiens clone 1
gb_hn1:CEY51A2_3	+	49.00	94.24	3.2e+03	102049 ! Continuation (4 of 4) of CEY
gb_hn1:CEY51A2D	+	49.00	91.41	4.6e+03	139259 ! AL021497 Caenorhabditis eleg
gb_hn1:CEY38F1	-	49.00	89.15	6.1e+03	178443 ! Z98861 Caenorhabditis eleg
gb_hn1:CEY068794	-	49.00	84.49	1.1e+04	297866 ! AC006794 Caenorhabditis eleg
gb_hn1:CEY068842	-	49.00	84.46	1.1e+04	299015 ! AC006842 Caenorhabditis eleg
gb_hn1:CEY06881	-	49.00	84.45	1.1e+04	299300 ! AC006881 Caenorhabditis eleg
gb_hn1:CEY06906	-	49.00	83.90	1.2e+04	317781 ! AC006906 Caenorhabditis eleg
gb_hn1:CEY46G5	-	49.00	83.73	1.2e+04	323720 ! Z92852 Caenorhabditis eleg
gb_hn1:CEY67H2	-	49.00	83.54	1.2e+04	330724 ! AL022475 Caenorhabditis eleg
gb_hn1:CEY12A2682	-	48.00	124.67	63.93	2556 ! AF142682 Ormosia formosana mat
gb_hn1:CEY12A2682	-	48.00	102.39	1.1e+03	29546 ! AF003390 Caenorhabditis eleg
gb_hn1:CEY12A2682	-	48.00	102.25	1.1e+03	30016 ! Z96101 Caenorhabditis eleg

gb\_in2:CEY12B5 - 48.00 101.53 1.2e+03 32488 ! AF100307 Caenorhabditis el  
 gb\_in1:CEW02B8 + 48.00 101.51 1.2e+03 32568 ! Z81136 Caenorhabditis eleg  
 gb\_in1:CEC34C12 - 48.00 100.91 1.3e+03 34759 ! Z46996 Caenorhabditis eleg

seq\_name: gb\_hn1:CEY38E10\_3

seq\_documentation\_block:

WPCOMMENT  
 Sequence split into 4 fragments LOCUS CEY38E10 Accession AL021149

Fragment Name Begin End  
 CEY38E10\_0 1 110000  
 CEY38E10\_1 100001 210000  
 CEY38E10\_2 200001 310000  
 CEY38E10\_3 300001 393707  
 Continuation (4 of 4) of CEY38E10 from base 300001 (AL021149 Caenorhabditis elegans c

## alignment\_scores:

Quality: 55.00 Length: 20  
 Ratio: 3.438 Gaps: 0  
 Percent Similarity: 80.000 Percent Identity: 60.000

## alignment\_block:

US-08-653-294-31 x CEY38E10\_3/rev ..

Align seg 1/1 to reverse of: CEY38E10\_3 from: 1 to: 93707

1 TyrArgLeuAlaIleArgLeuAsnGluArgTyrArgLeuAlaIleArgLe 17  
 ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :  
 15818 TATAGGCTTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCT 15769

17 uasnGluArg 20

15768 TAGCGGTAGG 15759

seq\_name: gb\_in1:CEY38E10A

## seq\_documentation\_block:

LOCUS CEY38E10A 110932 bp DNA INV 19-NOV-1999  
 Caenorhabditis elegans cosmid Y38E10A, complete sequence.

ACCESSION AL110484

VERSION AL110484.1 GI:5824723

KEYWORDS HTG.

SOURCE Caenorhabditis elegans.

ORGANISM

Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
 Rhabditina; Rhabditoidea; Rhabditidae; Pelodierinae; Caenorhabditis.

REFERENCE 1 (bases 1 to 110932)

AUTHORS

Wilson,R., Ainscough,R., Anderson,K., Baynes,C., Berks,M.,  
 Bonfield,J., Burton,J., Connell,M., Copsey,T., Cooper,J.,  
 Coulson,A., Craxton,M., Dear,S., Du,Z., Durbin,R., Favello,A.,  
 Fulton,L., Gardner,A., Green,P., Hawkins,T., Hillier,L., Jier,M.,  
 Johnston,L., Jones,M., Kershaw,J., Kirsten,J., Laister,N.,  
 Latreille,P., Lightning,J., Lloyd,C., McMurray,A., Mortimore,B.,  
 O'Callaghan,M., Parsons,J., Percy,C., Rifken,L., Roopra,A.,  
 Saunders,D., Showkneen,R., Smaildon,N., Smith,A., Sonnenhammer,E.,  
 Staden,R., Sulston,J., Thierry-Mieg,J., Thomas,K., Vaudin,M.,  
 Vaughan,K., Waterston,R., Watson,A., Weinstock,L.,  
 Wilkison-Sproat,J. and Wohlman,P.

2.2 Mb of contiguous nucleotide sequence from chromosome III of C.

elegans

Nature 368 (6466), 32-38 (1994)

94150718

2 (bases 1 to 110932)

WALLIS,J.

Direct Submission

Submitted (19-NOV-1999) Louis, MO 63110, USA. E-mail:

jesganger.ac.uk or rw@nematoe.wustl.edu

Coding sequences below are predicted from computer analysis, using

predictions from Genefinder (P. Green, U. Washington), and other

available information.

For a graphical representation of this sequence and its analysis

see:..

http://webace.sanger.ac.uk/cgi-bin/display?db=wormaceclass=Sequence subject=y38E10A  
 current sequence finishing criteria for the C. elegans genome sequencing consortium are that all bases are either sequenced unambiguously on both strands, or on a single strand with both a dye primer and dye terminator reaction, from distinct subclones. Exceptions are indicated by an explicit note.  
 IMPORTANT: This sequence is NOT necessarily the entire insert of the specified clone. It may be shorter because we only sequence overlapping sections once, or longer because we arrange for a small overlap between neighbouring submissions.  
 IMPORTANT: This sequence is not the entire insert of clone y38E10A. It may be shorter because we only sequence overlapping sections once, or longer because we arrange for a small overlap between neighbouring submissions.  
 The true left end of clone Y46G5 is at 110833 in this sequence. The true right end of clone F49C5 is at 100 in this sequence. The start of this sequence (1..100) overlaps with the end of sequence Z81544. The end of this sequence (110833..110932) overlaps with the start of sequence A110485.

FEATURES  
 source  
 1..110932  
 /organism="Caenorhabditis elegans"  
 /db\_xref="taxon:6239"  
 /clone="y38E10A"  
 complement(321..746)  
 /gene="y38E10A.1"  
 complement(join(321..371,418..518,564..645,690..746))  
 /gene="y38E10A.1"  
 /note="predicted using GeneFinder"  
 /codon\_start=1  
 /protein\_id="CAB54392.1"  
 /db\_xref="GI:5824724"  
 /translation="MIPYFLLSTSAEANYFSSDPFQFSTSCDPLETTTVAETSTPYHPMNDKDFTHSLTQEQSFNMLYTTGLARAKFCGSLPAKIN"  
 complement(1814..3652)  
 /gene="y38E10A.2"  
 complement(join(1814..1921,1966..2097,2146..2247,2618..2746,2791..2901,3387..3518,3563..3652))  
 /gene="y38E10A.2"  
 /note="predicted using GeneFinder"  
 /codon\_start=1  
 /protein\_id="CAB54393.1"  
 /db\_xref="GI:5824725"  
 /translation="MANFKTEFGISSQISWYAMLYNNYALKPILFAGNDSTADYSGETWRDACYHRFTAQPDALYVAIWLVDNDATTTTPMAWNFECTYPMSTQASINTLKHYNFALKPILFGGNDSTADYSGETWRDPCYHRFYAQNATYVIYQVENDATPLA WNFRTLPFGIGTOOTWDYKFOYDNFALKPILFAGNDSTADYSGETWRDPCYHRLYQ PNATFVYQVENDMYCEVMRRTCCLTAVTPSFARYLTRVEMRGDGCAL"  
 complement(4160..4891)  
 /gene="y38E10A.3"  
 complement(4160..4891)  
 /gene="y38E10A.3"  
 /note="predicted using GeneFinder"  
 /codon\_start=1  
 /protein\_id="CAB54394.1"  
 /db\_xref="GI:5824726"  
 /translation="MSELEYLVKLLVYASQVQAEQKLRKQKEKDALIKTTQVKLEVEA LGNPAPNRNSEDISKREIQSLKQVTVLKEEIGLEGSKAIAVANQEQKLDLADKEK ECKRFTVLYQNIQLQDPSRDQVADYLFQFRTVYKLEVRQKEEKLMAENKY NYFNCAEKCRQIETQSAARREDLLRDKMTLVKNNAKSEELLGQLDRDSTERCEEM EIDNCIOLLEQKREHLLNYYASLG"  
 complement(5436..7845)  
 /gene="y38E10A.4"  
 complement(join(5436..5499,5588..5715,5919..6616,6689..6820,6870..7011,7645..7845))  
 /gene="y38E10A.4"  
 /note="predicted using GeneFinder"  
 /codon\_start=1  
 /protein\_id="CAB54395.1"  
 /db\_xref="GI:5824727"  
 /translation="MQLFOSIFLILFSHAQLSPVCNNGYTLVNNKRLRFLKPTET HKTAELTCLNGATLANIKSIDNRATITFVGGASNSIWIGLFTKSSAKECFWDDD

SSGADQFQNEKSGPLVEKGRVYSQVAFDRGOWSSGDCCKESRAVFCELPVTNEDQ  
 CQXNYGCFYDFAPLFFVSAQKICKENCGLIISILSPMENRYINAFYFYSALLIGA  
 TWSNYSYTFDSSWSYHNIDHTARRGGVCLAMSTGTSMWPTGSPYVDCASNSF  
 LCRPAGTSCPWEPTPTTPVSPMSNMIMAPGTISSPAYGYDNNMYCYLLS  
 TGAYNILLKFTSTNLNDYTVYDGETTSSPMLGFSGCFNETPINLSTGMTLKR  
 GHKIMTKIIICTINIPVYGKMKLAYKPSWSMYQVGLCIKPPIDLPASLLRPIH  
 VIPA"  
 complement(8541..11513)  
 /gene="y38E10A.5"  
 complement(join(8541..9340,9391..9519,10537..10678,  
 10971..11187,11410..11513))  
 /gene="y38E10A.5"  
 /note="predicted using GeneFinder"  
 /codon\_start=1  
 /protein\_id="CAB54396.1"  
 /db\_xref="GI:5824728"  
 /translation="MIPQYKSLKIQYFLKARSWRTGTVVVLKLRNFGHLIMQFYSV  
 LLVSLLLLTISEASYPVCTSGTGLINGKCLIRIFVDVSTHTAAEKTCKGYGATLVY  
 KKSIDNRATADTGTNNANLFWGLYCFDSVSKLWDDATGSAEVYDNFAAGPHTA  
 LGNCVYSVQALAGMWSLSDCNDRRSFICELPASHADPCPYNNGFCYFHTASTY  
 LKQKICEQCGNLASIHSENRYIMTFQGRATKEDLLGLGMPADDDYVNWVDSLM  
 EYENFDPINVRDSVCVIMSNDRSPIALGMWYSGECKNEYSVVCKPAGTCCPNPTI  
 VPYTPPAVQSCNSYMAPSEITSPHPYNYNSDFCSYQI1STLSYNYLLRFSFDF  
 TEKVNVDVTVYDGDSTNDPVIGYAGSYFPTVISTGNKMLVTFKTDKSNTRQGSGR  
 ITSYSTTESGERD"  
 complement(12499..25367)  
 /gene="y38E10A.6"  
 complement(join(12499..13582,13577..13651,14625..15155,  
 16336..16509,17362..18123,19159..19332,20696..21019,  
 22013..22186,23133..23294,23443..23562,24374..24508,  
 25200..25367))  
 /gene="y38E10A.6"  
 /codon\_start=1  
 /protein\_id="CAB60334.1"  
 /db\_xref="GI:6425293"  
 /translation="MMFKRRDAVYAEHHGYELPCQMLLDEAYNKCRIISDKDRR  
 VLKAKVEERTLEEFFOKNACQDQIEMAKINRPTVFLSKWFSKARRSGYVKRQDP  
 EPGIHLPEYQKHQVQDGAKYELAEIAGVNPNTYKSWYMRRLRAQKADHGEEL  
 PQCKMELQEFAGHRKLSKGFQYIGDKILGTANFVSNYTFIRKMSKGAKRGERLSD  
 SDDDDPEDDGEEKNEEKNEEAKAAKAAQKMEAMTEAMTEERQYQDNVLY  
 TIFQYFVDVDRYQISETGTPAEIIRDMFQKRVSSVEEYKDGKELPKQMKLLH  
 ESYERCPMLGDDVRYNLAKTFLPPKISITNFFINRNRKTAIKRAESSVVDGAEPN  
 NDGFENLEAPILKLSURKVPRIIGDFSGVPEKNGSETAEKAOILKKNSETAE  
 IAKAQISEKPNVRLVDEVLSPQSEKIEPEKAOIEIEESPKEKAPKQPI  
 QEIPKSEKAPAPQIEKSPKSEKQEQEIQPKSEKSEKPEIEELPTFKSSA  
 PAQTEPESIDEIFAQTESLLNAVSEFQVDDRTNVLQSKQIRIGMPQIREFRKR  
 ESVEEHRITGELPKQMLLHAYQRCPLLEDARDLDVEKTLQLPKWYNTFINRS  
 RKQAANAENAEPTSDATTSDDGFDENIENFVALTOTRKSARPAKPYDDFGE  
 DADLLELRATTNPPAAVATAFTIGSHIIISPSOKHKSTQTTDFGKPEAKPEAPK  
 AVVTELPDPTDEEFVADLFEPLNDDFEPRDKRPAGPIVIRPPPKQPKLSIKTR  
 ILSEAYANLLKDKLTAHEIATFANRNNTEIDQPRIHVPKKE"  
 complement(37113..36098)  
 /gene="y38E10A.7"  
 join(36098..36285,36334..36807,36859..36979,37036..37113)  
 /gene="y38E10A.7"  
 /note="predicted using GeneFinder"  
 /codon\_start=1  
 /protein\_id="CAB54398.1"  
 /db\_xref="GI:5824730"  
 /translation="MAVSTVLLVLAALLSTADSPSFRAFQIOTYGSATLNLARTD  
 LGAGSVGGGLHDLGSLPTTRTVILVHGTNTSAGTAPQAAFYFRANGSSEETVATY  
 GDAGVTATNVKMLCEVQOIRNMIAVNAFTQKVNIGYSMGSPTRAKAILGKGA  
 ENYVQLGAPLTSIETIYISVAGNRTSLCDILEAPLVPTCNTKNGKLCSSDELTDI  
 RSVAYEGQIIFSIYGFSDDKVGFNVCGRVSRIDGATAEKDNVPGNHDALIANVGA  
 TADLLRNHAF"  
 complement(37322..38800)  
 /gene="y38E10A.8"  
 join(37322..37479,37532..37768,37813..37926,37979..38207,  
 38714..38800)  
 /gene="y38E10A.8"  
 /note="predicted using GeneFinder"  
 /codon\_start=1



available information.  
For a graphical representation of this sequence and its analysis see:  
<http://webcam.sanger.ac.uk/cgi-bin/display?db=wormacsc&sequence=object-Y41C4A>  
Current sequence finishing criteria for the C. elegans genome sequencing consortium are that all bases are either sequenced unambiguously on both strands, or on a single strand with both a dye primer and dye terminator reaction, from distinct subclones. Exceptions are indicated by an explicit note.  
**IMPORTANT:** This sequence is NOT necessarily the entire insert of the specified clone. It may be shorter because we only sequence overlapping sections once, or longer because we arrange for a small overlap between neighbouring submissions.  
**IMPORTANT:** This sequence is not the entire insert of clone Y41C4. It may be shorter because we only sequence overlapping sections once, or longer because we arrange for a small overlap between neighbouring submissions.  
The true left end of clone Y41C4 is at 1 in this sequence. The true left end of clone C24H11 is at 128957 in this sequence. The true right end of clone Y66A7 is at 107570 in this sequence. The start of this sequence (1..91) overlaps with the end of sequence AL032622.

right end of clone Y6A7 is at 120570 in this sequence. The start of this sequence (1..91) overlaps with the end of sequence AL032622.

The end of this sequence (128957..129060) overlaps with the start of sequence Z81475.

Location/Qualifiers

Organism= Caenorhabditis elegans  
/db\_xref="taxon:6239"

37715. 55607  
/C10ne- I41C4A

37715. 55607  
/gene="I4IC4A.4a"

/gene=Y4IC4A.4B*
to1n(37715)  37786.38474  38550.4

```
33355. :33361,34534. :54/21,55488.
/gene="Y41C4A.4b"
```

```
/couon_start=1
/protein_id="CAB54382.1"
```

```
/do_xref="GI:58247356"
/translation="MATASTNCSPVATSPLMWLLFKALQEGGSEDEARRRRLQNL
RRPSYRMILKDLTADKYMKKEPTPPSSVDASPLQFQSVMRPPTAPPTSAATPNR
ILPSSNAASPXGSPGLGSSILSNQPLVLPFAPINGDQDFSAIAAASQPKGLGWGFSG
```

GGFQNGLGSGGGVPGFSSGIAGMSVQFPIS  
PMLGIPAVSEFEGTTDWOSPM LSGGVSSSPST

VRLKNREAAKRCRRKKREYVRCLENRVSVLENC  
101p(37715) 37786.38474 38550.4

53355. .53561,54534. .54727,55488:  
/gene="Y41C4A.4b"

/note="similar to bZIP transcription factor 3 comes from this gene: CD"

from this gene; cDNA EST yk624a3.3

```
/codon_start=1
```

```

/db_xref="GI:5824755"

```

RRPSYRMILKDLETADKVMKKEPEETPPSSVDAS

PQNGLGCGGGGVPSSGIAGMSVQPPST

LLKNREAAKECRKKKEYVVKLENRVSVLENQNK

```
45015. :45503
/gene="Y41C4A.3"
```

43013. .43303  
/gene="Y41C4A.3"

```
/note="predicted using Generalizer"  
/codon start=1
```

```
/protein_id="CAA21546.1"
/db xref="GI:3880899"
```

```
/translation="MGQFII EYKIRIVSQTHLE  
PLIDI STKKKIL ARASEGATSSSSSPIPPASTATL
```

GFTPLPAQAMMANFPPNALYQLGTPTFFENGNNR





MEDLINE  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

94150718  
2. (bases 1 to 139259)  
McMurray, A.  
Direct Submission  
Submitted (23-JAN-1998) Louis, MO 63110, USA. E-mail:  
jes@sanger.ac.uk or rv@nematoe.wustl.edu  
Coding sequences below are predicted from computer analysis, using  
predictions from Genefinder (P. Green, U. Washington), and other  
available information.  
For a graphical representation of this sequence and its analysis  
see:-  
http://webace.sanger.ac.uk/cgi-  
bin/displaydb-wormaceclass-sequence.sobject-Y51A2D  
Current sequence finishing criteria for the C. elegans genome  
sequencing consortium are that all bases are either sequenced  
unambiguously on both strands, or on a single strand with both a  
dye primer and dye terminator reaction, from distinct subclones.  
Exceptions are indicated by an explicit note.  
IMPORTANT: This sequence is NOT necessarily the entire insert of  
the specified clone. It may be shorter because we only sequence  
overlapping sections once, or longer because we arrange for a small  
overlap between neighbouring submissions.  
IMPORTANT: This sequence is not the entire insert of clone Y51A2D.  
It may be shorter because we only sequence overlapping sections  
once, or longer because we arrange for a small overlap between  
neighbouring submissions.  
The true left end of clone Y69H2 is at 139160 in this sequence. The  
true right end of clone ZK228 is at 104 in this sequence. The true  
right end of clone Y37H2 is at 28891 in this sequence. The start of  
this sequence (1..104) overlaps with the end of sequence Z82086.  
The end of this sequence (139160..139259) overlaps with the start  
of sequence Z98877.  
Location/Qualifiers  
1..139259  
/organism="Caenorhabditis elegans"  
/db\_xref="taxon:6239"  
/chromosome="v"  
/clone="Y51A2D"  
complement(2063..7967)  
/gene="Y51A2D.18"  
complement(join(2063..2272,2845..2970,3017..3084,  
3268..3494,4315..4411,4735..4880,5226..5333,5761..6002,  
6063..6205,7760..7967))  
/gene="Y51A2D.18"  
/note="similar to Sugar (and other) transporter; cDNA EST  
yk30297.3 comes from this gene"  
/codon\_start=1  
/protein\_id="CAAL16411.1"  
/db\_xref="GI:3881048"  
/translation="MSSTSSNSEDSTINLPENGAKLEFLEKEPEQVHGLDDYIQMGKY  
VLVGLAAEILLPOVSMFYMYAGAPRIVSDINIVDSLEGEICFELDKMNC  
SHPEFQQRISINVEINFCNTSKLVKNSISVOMFGVLGTSVFGQISDSFGKISQ  
VASLGMGLIVVQSNLHFTISRTIVGFTGGSSIIINVFIMENPKKHKRWIM  
AIVSPNMPITYFAWLASDWKSLAINAFMPCIPGILFFOFFIHESPRWLVTKWIS  
AVOVSPNPSKSKNEIFLLSVCAEIGNHNCAGHCATSIINYGVLFNKEKISG  
SIWNSVTVGLMYACNLSPGYADLFKSGIKRFTISGLVIVLSLVVGYVALHL  
NHEMKVIRISILLASMTSQIYADGIVSAELEFPPIFTIGYAFQLNRCVGVLSF  
FIYLDYWLALPFCFMIPLFLSIDTFTSFCELLPDKGRHLVHEMPHQQWFGAGREE  
IVEEPPDTLEAKIEA"  
15721..28788  
/gene="Y51A2D.19"  
join(15721..15735,15871..15989,16146..16248,16410..16635,  
16844..17166,18096..18404,19389..19697,20436..20548,  
21854..22160,25686..25985,26726..27008,27230..27360,  
27740..27934,28423..28584,28726..28788)  
/gene="Y51A2D.19"  
/codon\_start=1  
/protein\_id="CAB54459.1"  
/db\_xref="GI:5824814"  
/translation="MKFFVAKDWAGELISQSITGRFLVLLVILSILGSLIIYFYDAS  
FQNFQVETCPWDSQSDLDLGNIFLFFIRFAASDKWFLLEMLFLEMLFFI  
PPSFVAYLQRMGLMGFWTHIRTNLNGFFIFPKNRTPDVLSFNLPISAFIDFIR  
ALRLMTVPDIQVLNLTSSIRLTQLTVIFVAVCLTGAGLVHLLNSGSDFFKFGIN

PHRTVADSVYFVLVTMSTGYGDIYCTTLCGRLEMIFFILFGLAMFASGVVPEIADLI  
GNQKYGEGVKGEGKKHIVCGHIIYDSVSHFLODFLEDHEDDDVDEVEVLRHVYVD  
LELEGLKRFHTKYEFFTGTVDSDLDLVRKIGDADACILANKYSINTPDAEDANIM  
RVISIKNYSDDIRVIVOLMGYHNKATLLNIPSWMRGDDVCLAEKLGFTAQSCILA  
PGFTMMANLFAWSEKTPHTPLNLDYLRGDMGEMYTESSPANNFSPFAEALL  
FNRGLILLALIKDENKECIAINPGPHIVIQPOTGFFTAQDEYKRAFFFWCKQ  
CHDDIKDVLKCKCKCKYQIFSEKIEPKLKRITQFQDMKYDSTGTFWCP  
SNLEDCVLERHOAAVTLNGHVHVCLEFADQDPLGLRNLIMPSSNFHVELKHV  
VVGDEYLRKEKWTLYNPKISILNGSLRADLRAYNIMCDMKVISAARPNTED  
TLADKEAILASLNKAMQFDDTLGFFPMRHTGDRSPIGSPISMKVKKKAGFTNPM  
ITELVGTGATPELEILAEAGLGGYSTPETLSNRDRIQIISLQDNPYDFNSALT  
TVGAMETIARYGOLGICGLYLRHDQDPSMKRYVITNPPAELRIKNTDYVYVLEOF  
DGLVEYEGKRHE"  
29727..33453  
/gene="Y51A2D.1"  
join(29727..29875,29929..30009,30982..31324,32223..32517,  
33173..33453)  
/gene="Y51A2D.1"  
/note="similar to Cysteine proteases (2 domains): cDNA EST  
EMBL:D66314 comes from this gene; cDNA EST EMBL:D70053  
comes from this gene"  
/codon\_start=1  
/protein\_id="CAAL16404.1"  
/db\_xref="GI:3881041"  
/db\_xref="SPTREMBL:G62484"  
/translation="MLFLOFFLNFQALGIFALESSSEPEFEINIDRHPKVKYQEF  
ELFKAFSTIKSEAKENQLRQNFVSRNNVRLNKAOKAGNSFVAVNQSDLTIS  
VHQKRSFPPNLTENSEVFKNFKLLGTRKTRKNSFARNFDRSQVGRVIVGP  
INQKQACCGWFAVMTLETIYAVNGVFKLSHPALPNFSDFFFEFLAKLN  
MFLSDQEMCDCATDTRAGACGGLMGVYAINNGLASEFDYDFEQNRATRGIT  
CEAMDDKTFFPPVNFAGTAFLQYKSGVLVTEDCDLAGTVHAGATVVGGEENDLRGR  
SOREWIMKNSGVSGWGTGYKLVIRGNKWCIGERGAIGANMEEHYHG"  
37734..44550  
/gene="Y51A2D.4"  
join(37734..37745,38371..38584,39138..39341,40053..40300,  
41019..41539,42067..42203,43039..43388,44416..44550)  
/gene="Y51A2D.4"  
/note="predicted using Genefinder; similar to Sugar (and  
other) transporters"  
/codon\_start=1  
/protein\_id="CAAL16400.1"  
/db\_xref="GI:3881037"  
/translation="WVAFAFSFGODKPAHTPKLGLFVILAAASVIGGFLGYDTS  
VYSAMLYPDAPLGKPMPTVQGEIVLSIPGMAVGLSMGTSSDVGIRKRVILGAS  
AFTGALYCAASVYKIMLLYGRVLLGIAIGFASMLVPLVLETATHVGRMLVAF  
LMSFGVAVANTIGTGAFTIDIPNGWRMLFAFAVPSIIQFCVNFLEPRLWYEN  
GFEETREVLKYNIGDEWEYEMAEIIAFEDQAKENAKHASGPIWRILKTPHY  
LKACFTIGMLQAFQOLAGINTILYITADIRSSGINSNHTIIVISLLCNFPGFY  
PKSLIEKVRRIIFLFSGLVSLVFIGVAFLLVNHDSAAITPANQSGTNSFNSTPD  
AKGCMAYSNCDVTTDAGCFCHDANTKOGYCLPAGFDNPEVTSYSGTNSNGSIAN  
NFKWXYCDTYTLPIIACGYLLTFSSGFTSLPWLNSSEYPPHARSTCAVISTT  
SNWFNLIITLITSLTVIGYKFAWLFAGLTIVIAFIFLFLVPTKGTGSIIEVEML  
FNKKORRAESRRRETTVEVSRMNSTVTSFGOHNEVHY"  
45748..49734  
/gene="Y51A2D.5"  
join(45748..45976,46209..46412,46825..47178,47593..48174,  
48739..49079,49603..49734)  
/gene="Y51A2D.5"  
/note="predicted using Genefinder; similar to Sugar (and  
other) transporters; cDNA EST yk308e4.3 comes from this  
gene; cDNA EST EMBL:C12832 comes from this gene; cDNA EST  
yk234f1.5 comes from this gene; cDNA EST yk240d11.3 comes  
from this gene; cDNA EST yk240d11.5 comes from this gene;  
cDNA EST yk43b6.5 comes from this gene; cDNA EST  
yk305g3.5 comes from this gene; cDNA EST yk308e4.5 comes  
from this gene; cDNA EST yk318a9.5 comes from this gene"  
/codon\_start=1  
/protein\_id="CAAL16405.1"  
/db\_xref="GI:3881042"  
/translation="WYAVEKYSESRPERKPNKLGFFVYLLGSAALIGGLFGYDT  
SVYSAMLYPDAPLGKPMGTWKEVIVSITPGMAAVGAWFGAGSDRGKRFILGS  
TLIFVCGAVICAVANTKIYMLIGRIFLGVIGFASVMPVPLGEASPTHVRGLTVSAF

FEATURES  
source

gene  
CDS

gene  
CDS

AMMSFGQVANNIMGVSFYWEPTTIGWLMFAFAGIPALIQVCFIPLPPTRWLYE  
 NGHTEOAQVLEKIYGGTEWEYELAIKTYAEROKMEERKSGPVIVRIKLTPH  
 VLKACFIGSLQAQFOLAGITLYVTADILRSAGIENHVIINISVILSICNLGPE  
 AKPMFIEKLRKLEFLFCAGGVVSLVLIGVSLVGVNDSPNDRSAYLAGNYQSN  
 GEASCLMSLSCVTSSEHCGEDSETRTGFLCPVDHNDVLYSGCLCTNGLDKS  
 NSSFNATSYVQKHCHTSTILPVMGVYLLTFSCGFTSLPVLNSEFPYMWARS  
 TCVSISTLSNVFNLIATILVLSLTHATKYGAFLYAIFTIATFIYFLVLPETTY  
 SIDEVEMLFMKNRQRIANQARQAKDAADKDKNSSTSLSTETITM  
 58395 69879  
 /gene="Y51A2D.7a"  
 58395 69879  
 /gene="Y51A2D.7b"  
 join(58395..58445,58657..58792,59709..59891,61464..61787,  
 63584..63948,64698..64888,66270..66366,68189..68665,  
 69661..69879)  
 /gene="Y51A2D.7b"  
 /note="cDNA EST yk269a8.5 comes from this gene; cDNA EST  
 yk269a8.3 comes from this gene; cDNA EST yk332b9.3 comes  
 from this gene; cDNA EST yk332b9.5 comes from this gene;  
 cDNA EST yk348d3.3 comes from this gene; cDNA EST  
 yk348d3.5 comes from this gene; cDNA EST yk443f6.3 comes  
 from this gene; cDNA EST yk443f6.5 comes from this gene;  
 cDNA EST yk382c7.3 comes from this gene; cDNA EST  
 yk575e9.3 comes from this gene; cDNA EST yk619h9.3 comes  
 from this gene; cDNA EST yk673e8.3 comes from this gene"  
 /codon\_start=1  
 /protein\_id="CAA16406.1"  
 /db\_xref="GI:3881043"  
 /translation="MQRNPELQEAQSOLIVDSLDKPAAGRGSGVNRLLGFAFFKL  
 VTCSLKLQILVNTNCHLITPNVIRAIQVTVKSLILPAITFSDFIKQIVEDVDP

alignment\_scores:  
 Quality: 53.00 Length: 20  
 Ratio: 3.118 Gaps: 0  
 Percent Similarity: 85.000 Percent Identity: 50.000

alignment\_block:  
 US-08-653-294-31 x CEY51A2D/rev ..

Align seg 1/1 to reverse of: CEY51A2D from: 1 to: 139259

1 TyrArgLeuAlaIleArgLeuAsnGluArgTyrArgLeuAlaIleArgLe 17  
 137264 TTATAGCTTAGGCTAAGCTTAGACCTAGATTAGCTTAGGCTTAGGCT 137215

17 uAsnGluArg 20

|||||

137214 TAGACGTCGG 137205

seq\_name: gb\_htg2:AC006889

seq\_documentation\_block:  
 LOCUS AC006889 267118 bp DNA HTG 26-FEB-1999  
 DEFINITION Caenorhabditis elegans clone Y6584, \*\*\* SEQUENCING IN PROGRESS \*\*\*,  
 6 unordered pieces.

ACCESSION AC006889

VERSION AC006889.2 GI:4309908

KEYWORDS HTG; HTGS\_PHASE1.

SOURCE Caenorhabditis elegans.

ORGANISM Caenorhabditis elegans.

Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
 Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.

REFERENCE 1 (bases 1 to 267118)

AUTHORS Waterston,R.H.

TITLE The sequence of Caenorhabditis elegans clone

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 267118)

AUTHORS Waterston,R.H.

TITLE Direct Submission

JOURNAL Submitted (24-FEB-1999) Genome Sequencing Center, Washington

University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
 MO 63108, USA

COMMENT On Mar 1, 1999 this sequence version replaced gi:4263459.

\* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 6 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

\* 1 5875: contig of 5875 bp in length  
 \* 5876 5889: gap of unknown length  
 \* 5890 13618: contig of 7729 bp in length  
 \* 13619 13632: gap of unknown length  
 \* 13633 28840: contig of 15208 bp in length  
 \* 28841 28854: gap of unknown length  
 \* 28855 110971: contig of 82117 bp in length  
 \* 110972 110985: gap of unknown length  
 \* 110986 174804: contig of 63819 bp in length  
 \* 174805 174818: gap of unknown length  
 \* 174819 267118: contig of 92300 bp in length.

FEATURES Location/Qualifiers

1..267118  
 /organism="Caenorhabditis elegans"  
 /db\_xref="taxon:6239"  
 /clone="Y65B4"

BASE COUNT 83652 a 50557 c 48664 g 84175 t 70 others  
 ORIGIN

alignment\_scores:  
 Quality: 53.00 Length: 19  
 Ratio: 3.312 Gaps: 0  
 Percent Similarity: 84.211 Percent Identity: 57.895

alignment\_block:  
 US-08-653-294-31 x AC006889/rev ..

Align seg 1/1 to reverse of: AC006889 from: 1 to: 267118

2 ArgLeuAlaIleArgLeuAsnGluArgTyrArgLeuAlaIleArgLeuAs 18  
 134720 AGGCTTAGTATTAGGCTGAGCGCTAGATTAGGCTTAGGCTTAGGCTTAG 134671

18 nGluArg 20

|||||

134670 TGTTAGG 134664

seq\_name: gb\_htg2:AC006792

seq\_documentation\_block:  
 LOCUS AC006792 299202 bp DNA HTG 23-FEB-1999  
 DEFINITION Caenorhabditis elegans clone Y50C1a, \*\*\* SEQUENCING IN PROGRESS  
 \*\*\*, 10 unordered pieces.

ACCESSION AC006792

VERSION AC006792.1 GI:4263146

KEYWORDS HTG; HTGS\_PHASE1.

SOURCE Caenorhabditis elegans.

ORGANISM Caenorhabditis elegans.

Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
 Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.

REFERENCE 1 (bases 1 to 299202)

AUTHORS Waterston,R.H.

TITLE The sequence of Caenorhabditis elegans clone

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 299202)

AUTHORS Waterston,R.H.

TITLE Direct Submission

JOURNAL Submitted (23-FEB-1999) Genome Sequencing Center, Washington

University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
 MO 63108, USA

COMMENT \* NOTE: This is a 'working draft' sequence. It currently

\* consists of 10 contigs. The true order of the pieces

\* is not known and their order in this sequence record is

\* arbitrary. Gaps between the contigs are represented as

\* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

```

1 2371: contig of 2371 bp in length
2372 2385: gap of unknown length
2386 7309: contig of 4924 bp in length
7310 7323: gap of unknown length
7324 13580: contig of 6257 bp in length
13581 13594: gap of unknown length
13595 25198: contig of 11604 bp in length
25199 25212: gap of unknown length
25213 35793: contig of 10567 bp in length
35794 35780: gap of unknown length
35781 63108: contig of 27315 bp in length
63109 63122: gap of unknown length
63123 124971: contig of 61848 bp in length
124972 124984: gap of unknown length
124985 185093: contig of 60109 bp in length
185094 185107: gap of unknown length
185108 253768: contig of 68661 bp in length
253769 253782: gap of unknown length
253783 299202: contig of 45420 bp in length.

```

#### FEATURES

```

source
1. 299202
   /organism="Caenorhabditis elegans"
   /db_xref="taxon:6239"
   /clone="v50C1a"

```

BASE COUNT 95054 a 54073 c 54425 g 95524 t 126 others  
 ORIGIN

```

alignment_scores:
Quality: 53.00 Length: 19
Ratio: 3.312 Gaps: 0
Percent Similarity: 84.211 Percent Identity: 57.895

```

alignment\_block:  
 US-08-653-294-31 x AC006792

Align seg 1/1 to: AC006792 from: 1 to: 299202

```

2 ArgLeuAlaileArgLeuAsnGluArgTyrArgLeuAlaileArgLeuAs 18
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
173748 AGGCTTAGTATTAGGCTAGGCTAGATTAGGCTTATACITAGGCTTAG 173797

```

18 nGluArg 20

173798 TGTTAGG 173804

seq\_name: gb\_pr3:AC003070

```

seq_documentation_block:
LOCUS AC003070 129914 bp DNA PRI 10-FEB-1998
DEFINITION Homo sapiens chromosome 17, clone HCIT39G8, complete sequence.
ACCESSION AC003070
VERSION AC003070.1 GI:2865212
KEYWORDS HTG.
SOURCE human.

```

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 129914)  
 Birren,B., Fasnman,K., McKernan,K., Nusbaum,C., Richardson,P. and  
 Lander,E.

TITLE Homo sapiens chromosome 17, clone HCIT39G8

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 129914)

AUTHORS Birren,B., Fasnman,K., McKernan,K., Munro,C., Nusbaum,C.,  
 Richardson,P., Lander,E., Baldwin,J., Barna,N., Cantu,C., Chang,A.,  
 Cooke,P., Daly,M.J., Devon,K., Dewar,K., DuRette,B., Forrest,C.,  
 Gage,D., Gensheimer,S., Geraghty,K., Gilmartin,T., Hagos,B.,  
 Halphen,I., Harris,K., Howland,J.C., Huang,J., Hui,L., Jacotot,L.,

Kirby,A., Lane,M., Mackenzie,J., Marquis,N., McDermott,J.,  
 Molla,M., Morrow,J., Nachman,A., Naylor,J., O'Connor,T., Olotu,A.,  
 Peterson,K., Roberts,D., Rollins,G., Sarnaik,A., Shiue,P., Shyam,R.,  
 Stilwell,J., Stone,C., Strickland,C., Sydney,K., Tang,L.,  
 Zemtseva,I. and Zody,M.

#### TITLE

JOURNAL Direct Submission

#### REFERENCE

AUTHORS Submitted (06-NOV-1997) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 129914)  
 Birren,B., Fasnman,K., McKernan,K., Nusbaum,C., Richardson,P.,  
 Lander,E., Allen,N., Baker,J., Baldwin,J., Barna,N., Beckerly,R.,  
 Boutwell,C., Byrne,S., Cantu,C., Castle,A., Cooke,P., Daly,M.J.,  
 Depayre,E., Devon,K., Dewar,K., DuRette,B., Etenadi,S.,  
 Ferreira,P., Forrest,C., Gage,D., Gardyna,S., Gensheimer,S.,  
 Geraghty,K., Gilmartin,T., Gray,D., Hagos,B., Harris,K.,  
 Horton,L., Howland,J.C., Hui,L., Jacotot,L., Linton,L.,  
 Mackenzie,J., Marquis,N., McEwan,P., McGurk,A., Meidrim,J.,  
 Molla,M., Morris,W., Morrow,J., Nachman,A., Naylor,J., O'Connor,T.,  
 Pavlin,B., Peterson,K., Ranganath,S., Riley,R., Roberts,D.,  
 Rollins,G., Rossello,R., Roy,A., Shyam,R., Soohoo,S., Stilwell,J.,  
 Stone,C., Strickland,C., Sydney,K., Tang,L., Vassiliev,H., Vo,A.,  
 Wagner,A., Wheeler,J., Wu,Y., Ye,W.J., Zemtseva,I., Zhao,J. and  
 Zody,M.

#### TITLE

#### JOURNAL

#### COMMENT

Submitted (10-FEB-1998) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Feb 11, 1998 this sequence version replaced gi:2665502.  
 The Staden databases, finishing information, and all  
 chromatographic files used in the assembly of this clone are  
 available from our anonymous ftp site.

All repeats were identified using RepeatMasker: Smit, A.F.A. &  
 Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>.

#### FEATURES

##### source

```

1. 129914
   /organism="Homo sapiens"
   /db_xref="taxon:9606"
   /clone="HCIT39G8"
   /clone_lib="Research Genetics/Cal Tech CIB978SK-B (plates
1-194)"
   /map="17"

```

```

   /chromosome="17"
   complement(708..790)
   /rpt_family="LINE2"
   complement(1014..1172)
   /rpt_family="AluJo"
   complement(1173..1468)
   /rpt_family="AluSx"
   complement(1492..1614)
   /rpt_family="AluJo"
   complement(2534..2835)
   /rpt_family="AluSg"
   3577..3670
   /rpt_family="MIR"
   complement(5082..5209)
   /rpt_family="FLAM_A"
   5758..5859
   /rpt_family="MIR"
   /rpt_family="(CAA)n"
   complement(5870..6043)
   /rpt_family="MIR"
   6096..6165
   /rpt_family="MIR"
   6205..6262
   /rpt_family="AT-rich"
   9133..9161
   /rpt_family="GC-rich"
   complement(9521..9628)
   /rpt_family="(CGG)n"
   complement(10099..10183)
   /rpt_family="LINE2"
   10514..10783
   /rpt_family="AluJo"

```

```

repeat_region 11280..11348
/rpt_family="MIR"
complement(12323..12380)
/rpt_family="MER20"
repeat_region complement(12389..12715)
/rpt_family="THE1C"
repeat_region complement(12717..13010)
/rpt_family="AluSx"
repeat_region complement(13011..13051)
/rpt_family="THE1C"
repeat_region complement(13052..13209)
/rpt_family="MER20"
repeat_region complement(13218..13360)
/rpt_family="MIR"
repeat_region complement(13902..14031)
/rpt_family="MER63B"
repeat_region complement(14977..15110)
/rpt_family="FLAM_C"
repeat_region complement(15135..15438)
/rpt_family="AluSx"
repeat_region 15928..15992
/rpt_family="LINE2"
complement(15992..16091)
/rpt_family="LINE2"
repeat_region 16095..16315
/rpt_family="AluSg"
repeat_region 16407..16567
/rpt_family="LINE2"
repeat_region 16789..16938
/rpt_family="LINE2"
repeat_region 17242..17543
/rpt_family="AluSx"
repeat_region 17784..18081
/rpt_family="AluSx"
repeat_region 18676..18968
/rpt_family="AluSg"
complement(19585..19886)
/rpt_family="AluSx"
repeat_region complement(19911..20000)
/rpt_family="S8"
repeat_region complement(20334..20463)
/rpt_family="LINE2"
repeat_region complement(20562..20865)
/rpt_family="AluSx"
repeat_region complement(21117..21419)
/rpt_family="AluJo"
repeat_region complement(21773..22068)
/rpt_family="AluSg"
repeat_region complement(22732..22769)
/rpt_family="LINE2"
repeat_region complement(22778..23074)
/rpt_family="AluSg"
repeat_region 23104..23181
/rpt_family="MER64B"
repeat_region 23182..23289
/rpt_family="MER63B"
repeat_region complement(23312..23543)
/rpt_family="AluSg/x"
repeat_region 23544..23652
/rpt_family="MER63A"
repeat_region 23654..24050
/rpt_family="MER64B"
repeat_region complement(24051..24177)
/rpt_family="MIR"
repeat_region 24179..24309
/rpt_family="FLAM_C"
repeat_region 24336..24627
/rpt_family="AluSx"
repeat_region complement(24630..24664)
/rpt_family="MIR"
repeat_region 24691..24757
/rpt_family="MIR"
repeat_region 26560..26719

```

```

/rpt_family="LINE2"
complement(26913..27025)
/rpt_family="LINE2"
repeat_region complement(27148..27502)
/rpt_family="LTR16A"
repeat_region complement(27766..28067)
/rpt_family="AluJo"
repeat_region 28191..28324
/rpt_family="LINE2"
repeat_region 28553..28746
/rpt_family="MIR"
repeat_region complement(28867..28919)
/rpt_family="LINE2"
repeat_region 29355..29477
/rpt_family="(GGA)n"
repeat_region complement(29512..29643)
/rpt_family="MIR"
repeat_region complement(29714..29764)
/rpt_family="(CA)n"
repeat_region 29831..30134
/rpt_family="AluSx"
repeat_region complement(30143..30267)
/rpt_family="AluJo/FRAM"
repeat_region complement(31477..31502)
/rpt_family="AT-rich"
repeat_region 31507..31780
/rpt_family="AluJo"
repeat_region 31963..32036
/rpt_family="MIR"
repeat_region complement(32042..32344)
/rpt_family="AluSx"
repeat_region complement(32632..32771)
/rpt_family="MIR"
repeat_region complement(33178..33469)

```

alignment\_scores:  
 Quality: 52.00 Length: 18  
 Ratio: 3.714 Gaps: 0  
 Percent Similarity: 77.778 Percent Identity: 66.667

alignment\_block:  
 US-08-653-294-31 x AC003070/rev ..  
 Align seg 1/1 to reverse of: AC003070 from: 1 to: 129914

```

2 ArgLeuAlaIleArgLeuAsnGluArgTyrArgLeuAlaIleArgLeuAs 18
||||:||||| ||||||| ||||| ||||| ||||| ||||| |||||
120851 CGCGTGGCCATCGGCGGAATGAGCGCTACCGGCTGCTGCGGCGAGCAC 120802
18 nGlu 19
120801 CGAG 120798

```

seq\_name: gb\_in2:CELC45E1

seq\_documentation\_block:  
 LOCUS CELC45E1 27111 bp DNA INV 25-JUN-1998  
 DEFINITION Caenorhabditis elegans cosmid C45E1.  
 ACCESSION AF043694  
 VERSION AF043694.1 GI:2804421  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Caenorhabditis elegans  
 Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
 Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
 REFERENCE 1 (bases 1 to 27111)  
 AUTHORS  
 Wilson,R., Ainscough,R., Anderson,K., Baynes,C., Berks,M.,  
 Bonfield,J., Burton,J., Connell,M., Copsey,T., Cooper,J.,  
 Coulson,A., Craxton,M., Dear,S., Du,Z., Durbin,R., Favello,A.,  
 Fulton,L., Gardner,A., Green,P., Hawkins,T., Hillier,L., Jier,M.,  
 Johnston,L., Jones,M., Kershaw,J., Kirsten,J., Laister,N.,  
 Latreille,P., Lightning,J., Lloyd,C., McMurray,A., Mortimore,B.,  
 O'Callaghan,M., Parsons,J., Percy,C., Rifken,L., Roopra,A.,

Saunders, D., Showkeen, R., Smaldon, N., Smith, A., Sonhammer, E., Staden, R., Sulston, J., Thierry-Mieg, J., Thomas, K., Vaudin, M., Vaughan, K., Waterston, R., Watson, A., Weinstock, L., Wilkinson-Sproat, J., and Wohldman, P.  
2.2 Mb of contiguous nucleotide sequence from chromosome III of *C. elegans*  
Nature 368 (6466), 32-38 (1994)

JOURNAL  
MEDLINE  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

2 (bases 1 to 27111)  
Lewth, J., Graves, T., Sutterer, C. and Fronick, B.  
The sequence of *C. elegans* cosmid C45E1  
Unpublished (1998)  
3 (bases 1 to 27111)  
Waterston, R.  
Direct Submission  
Submitted (20-JAN-1998) Department of Genetics, Washington University, 444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
4 (bases 1 to 27111)  
Waterston, R.  
Direct Submission  
Submitted (25-JUN-1998) Department of Genetics, Washington University, 444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
Submitted by:  
Genome Sequencing Center  
Department of Genetics, Washington University,  
St. Louis, MO 63110, USA, and  
Sanger Centre, Hinxton Hall  
Cambridge CB10 1RQ, England  
e-mail: rwenematode.wustl.edu and jesse@sanger.ac.uk

NOTICE: This sequence may not be the entire insert of this clone. It may be shorter because we only sequence overlapping sections once, or longer because we provide a small overlap between neighboring submissions.

This sequence was finished as follows unless otherwise noted:  
all regions were double stranded or sequenced with an alternate chemistry; an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by a sequence from more than one subclone

#### NEIGHBORING COSMID INFORMATION:

The 5' cosmid is M01B12, 1300 bp overlap; 3' cosmid is K09H9, 200 bp overlap. Actual start of this cosmid is at base position 1 of CELC45E1; actual end is at 17470 of CELK09H9

#### NOTES:

Coding sequences below are predicted from computer analysis, using the program Genefinder (P. Green and L. Hillier, ms in preparation).

#### FEATURES

source  
1. .27111  
/organism="Caenorhabditis elegans"  
/strain="Bristol N2"  
/db\_xref="taxon:6239"  
/clone="C45E1"  
/chromosome="I"  
9241. .14524  
/gene="nhr-64"  
join(9241. .9514, 11160. .11234, 12520. .12715, 12857. .13015, 13204. .13326, 13511. .13617, 13670. .13725, 14276. .14524)  
/gene="nhr-64"  
/note="C45E1.1: coded for by *C. elegans* cDNA yk384e7.3;  
coded for by *C. elegans* cDNA yk384e7.3"  
/codon\_start=1  
/product="similar to the steroid/thyroid/retinoic nuclear hormone receptors; contains similarity to C4-type zinc fingers"  
/protein\_id="AAC24283.1"

/db\_xref="GI:2804422"  
/translation="MIFQNDHTTKIKIPKQFLSKFOLFPTHFPEFDSKTVLFSEEM  
TLERKEVSTSSQSPSSSENFVCAICGDRATKHYGAMSCDCKGFFRTIKRHR  
SVYRFGKQVDKAKRNSCRKCRFEDVCLRRKRDADVQTERDIRPANPLSNGSGG  
IYPPDLDTLRIRAEASTRTVITTAERKQATNDVTDSDNQTLTLMVEWAKY  
LEGFORVINTQVALLRHFSQAHLVMAAFRSIHLSDAVLTNETCLHKDKIPDMN  
RVAERIIDQVTPMRLSHMNEIEYIALKAIAFPDPLAGKITSETSDVEEMQRILS  
FERHYVSPYKDMPLRANLPLPMLAISRDVLVEDVQLAKLGLASIDNLMLELM  
LPNEKNMTDKTSPIMCHQ"  
complement(15723. .22627)  
/gene="C45E1.3"  
complement(15723. .22627)  
/gene="C45E1.4"  
complement(15723. .22627)  
18227. .18386, 20536. .20580, 21421. .21627, 22505. .22627))  
/gene="C45E1.4"  
complement(join(15723. .16044, 16879. .16957, 17048. .17134,  
18227. .18386, 20536. .20580, 21421. .21627, 22505. .22627))  
/gene="C45E1.4"  
evidence=not\_experimental  
/protein\_id="AAC24285.1"  
/db\_xref="GI:2804425"  
/translation="MSSLGTYFFLLVLLPVCATCYEEDYRPEGLTGHNGVFOALP  
WTKFELNLISSLHATANYPEVNLVRERMIISDIAPNDRRKIERMLKNLRPPVDEF  
LTEDETKVRKAHSERDVSILMNLNVLVIVVHSLKWEVKKIIGEHNFVDSN  
IRFTGLTDSKYIFLDLSLKHSHQSDLDSPSHCQCTDLKWEKFWSPSKELELQ  
KRKWCCTMEIIFEVKIYFSVKYKTFQVVCVQIKSFLSGKFDKFAAYPWTPI  
PGTHREKLLVCHSFVPLSAPRARMVTSVCGVLTLMVIVHEAAVWRAGEPTKEH  
VDGVIT"  
complement(join(15723. .16044, 16879. .16957, 17029. .17134,  
20536. .20580, 21421. .21627, 22505. .22627))  
/gene="C45E1.4"  
/note="C45E1.4"  
/codon\_start=1  
/protein\_id="AAC24284.1"  
/db\_xref="GI:2804424"  
/translation="MSSLGTYFFLLVLLPVCATCYEEDYRPEGLTGHNGVFOALP  
WTKFELNLISSLHATANYPEVNLVRERMIISDIAPNDRRKIERMLKNLRPPVDEF  
LTEDETKVRKAHSERDVSILMNLNVLVIVVHSLKWEVKKIIGEHNFVDSN  
IRFTGLTDSKYIFLDLSLKHSHQSDLDSPSHCQCTDLKWEKFWSPSKELELQ  
KRKWCCTMEIIFEVKIYFSVKYKTFQVVCVQIKSFLSGKFDKFAAYPWTPI  
PGTHREKLLVCHSFVPLSAPRARMVTSVCGVLTLMVIVHEAAVWRAGEPTKEH  
VDGVIT"  
complement(join(15723. .16044, 16879. .16957, 17029. .17134,  
20536. .20580, 21421. .21627, 22505. .22627))  
/gene="C45E1.4"  
/note="C45E1.4"  
/codon\_start=1  
/protein\_id="AAC24284.1"  
/db\_xref="GI:2804424"  
/translation="MSSLGTYFFLLVLLPVCATCYEEDYRPEGLTGHNGVFOALP  
WTKFELNLISSLHATANYPEVNLVRERMIISDIAPNDRRKIERMLKNLRPPVDEF  
LTEDETKVRKAHSERDVSILMNLNVLVIVVHSLKWEVKKIIGEHNFVDSN  
IRFTGLTDSKYIFLDLSLKHSHQSDLDSPSHCQCTDLKWEKFWSPSKELELQ  
KRKWCCTMEIIFEVKIYFSVKYKTFQVVCVQIKSFLSGKFDKFAAYPWTPI  
PGTHREKLLVCHSFVPLSAPRARMVTSVCGVLTLMVIVHEAAVWRAGEPTKEH  
VDGVIT"  
complement(join(15723. .16044, 16879. .16957, 17029. .17134,  
20536. .20580, 21421. .21627, 22505. .22627))  
/gene="C45E1.4"  
/note="C45E1.4"  
/codon\_start=1  
/protein\_id="AAC24284.1"  
/db\_xref="GI:2804424"  
/translation="MSSLGTYFFLLVLLPVCATCYEEDYRPEGLTGHNGVFOALP  
WTKFELNLISSLHATANYPEVNLVRERMIISDIAPNDRRKIERMLKNLRPPVDEF  
LTEDETKVRKAHSERDVSILMNLNVLVIVVHSLKWEVKKIIGEHNFVDSN  
IRFTGLTDSKYIFLDLSLKHSHQSDLDSPSHCQCTDLKWEKFWSPSKELELQ  
KRKWCCTMEIIFEVKIYFSVKYKTFQVVCVQIKSFLSGKFDKFAAYPWTPI  
PGTHREKLLVCHSFVPLSAPRARMVTSVCGVLTLMVIVHEAAVWRAGEPTKEH  
VDGVIT"  
complement(join(15723. .16044, 16879. .16957, 17029. .17134,  
20536. .20580, 21421. .21627, 22505. .22627))  
/gene="C45E1.4"  
/note="C45E1.4"  
/codon\_start=1  
/protein\_id="AAC24284.1"  
/db\_xref="GI:2804424"  
/translation="MSSLGTYFFLLVLLPVCATCYEEDYRPEGLTGHNGVFOALP  
WTKFELNLISSLHATANYPEVNLVRERMIISDIAPNDRRKIERMLKNLRPPVDEF  
LTEDETKVRKAHSERDVSILMNLNVLVIVVHSLKWEVKKIIGEHNFVDSN  
IRFTGLTDSKYIFLDLSLKHSHQSDLDSPSHCQCTDLKWEKFWSPSKELELQ  
KRKWCCTMEIIFEVKIYFSVKYKTFQVVCVQIKSFLSGKFDKFAAYPWTPI  
PGTHREKLLVCHSFVPLSAPRARMVTSVCGVLTLMVIVHEAAVWRAGEPTKEH  
VDGVIT"  
complement(join(15723. .16044, 16879. .16957, 17029. .17134,  
20536. .20580, 21421. .21627, 22505. .22627))  
/gene="C45E1.4"  
/note="C45E1.4"  
/codon\_start=1  
/protein\_id="AAC24284.1"  
/db\_xref="GI:2804424"  
/translation="MSSLGTYFFLLVLLPVCATCYEEDYRPEGLTGHNGVFOALP  
WTKFELNLISSLHATANYPEVNLVRERMIISDIAPNDRRKIERMLKNLRPPVDEF  
LTEDETKVRKAHSERDVSILMNLNVLVIVVHSLKWEVKKIIGEHNFVDSN  
IRFTGLTDSKYIFLDLSLKHSHQSDLDSPSHCQCTDLKWEKFWSPSKELELQ  
KRKWCCTMEIIFEVKIYFSVKYKTFQVVCVQIKSFLSGKFDKFAAYPWTPI  
PGTHREKLLVCHSFVPLSAPRARMVTSVCGVLTLMVIVHEAAVWRAGEPTKEH  
VDGVIT"  
complement(join(15723. .16044, 16879. .16957, 17029. .17134,  
20536. .20580, 21421. .21627, 22505. .22627))  
/gene="C45E1.4"  
/note="C45E1.4"  
/codon\_start=1  
/protein\_id="AAC24284.1"  
/db\_xref="GI:2804424"  
/translation="MSSLGTYFFLLVLLPVCATCYEEDYRPEGLTGHNGVFOALP  
WTKFELNLISSLHATANYPEVNLVRERMIISDIAPNDRRKIERMLKNLRPPVDEF  
LTEDETKVRKAHSERDVSILMNLNVLVIVVHSLKWEVKKIIGEHNFVDSN  
IRFTGLTDSKYIFLDLSLKHSHQSDLDSPSHCQCTDLKWEKFWSPSKELELQ  
KRKWCCTMEIIFEVKIYFSVKYKTFQVVCVQIKSFLSGKFDKFAAYPWTPI  
PGTHREKLLVCHSFVPLSAPRARMVTSVCGVLTLMVIVHEAAVWRAGEPTKEH  
VDGVIT"  
complement(join(15723. .16044, 16879. .16957, 17029. .17134,  
20536. .20580, 21421. .21627, 22505. .22627))  
/gene="C45E1.4"  
/note="C45E1.4"  
/codon\_start=1  
/protein\_id="AAC24284.1"  
/db\_xref="GI:2804424"  
/translation="MSSLGTYFFLLVLLPVCATCYEEDYRPEGLTGHNGVFOALP  
WTKFELNLISSLHATANYPEVNLVRERMIISDIAPNDRRKIERMLKNLRPPVDEF  
LTEDETKVRKAHSERDVSILMNLNVLVIVVHSLKWEVKKIIGEHNFVDSN  
IRFTGLTDSKYIFLDLSLKHSHQSDLDSPSHCQCTDLKWEKFWSPSKELELQ  
KRKWCCTMEIIFEVKIYFSVKYKTFQVVCVQIKSFLSGKFDKFAAYPWTPI  
PGTHREKLLVCHSFVPLSAPRARMVTSVCGVLTLMVIVHEAAVWRAGEPTKEH  
VDGVIT"  
complement(join(15723. .16044, 16879. .16957, 17029. .17134,  
20536. .20580, 21421. .21627, 22505. .22627))  
/gene="C45E1.4"  
/note="C45E1.4"  
/codon\_start=1  
/protein\_id="AAC24284.1"  
/db\_xref="GI:2804424"  
/translation="MSSLGTYFFLLVLLPVCATCYEEDYRPEGLTGHNGVFOALP  
WTKFELNLISSLHATANYPEVNLVRERMIISDIAPNDRRKIERMLKNLRPPVDEF  
LTEDETKVRKAHSERDVSILMNLNVLVIVVHSLKWEVKKIIGEHNFVDSN  
IRFTGLTDSKYIFLDLSLKHSHQSDLDSPSHCQCTDLKWEKFWSPSKELELQ  
KRKWCCTMEIIFEVKIYFSVKYKTFQVVCVQIKSFLSGKFDKFAAYPWTPI  
PGTHREKLLVCHSFVPLSAPRARMVTSVCGVLTLMVIVHEAAVWRAGEPTKEH  
VDGVIT"  
complement(join(15723. .16044, 16879. .16957, 17029. .17134,  
20536. .20580, 21421. .21627, 22505. .22627))  
/gene="C45E1.4"  
/note="C45E1.4"  
/codon\_start=1  
/protein\_id="AAC24284.1"  
/db\_xref="GI:2804424"  
/translation="MSSLGTYFFLLVLLPVCATCYEEDYRPEGLTGHNGVFOALP  
WTKFELNLISSLHATANYPEVNLVRERMIISDIAPNDRRKIERMLKNLRPPVDEF  
LTEDETKVRKAHSERDVSILMNLNVLVIVVHSLKWEVKKIIGEHNFVDSN  
IRFTGLTDSKYIFLDLSLKHSHQSDLDSPSHCQCTDLKWEKFWSPSKELELQ  
KRKWCCTMEIIFEVKIYFSVKYKTFQVVCVQIKSFLSGKFDKFAAYPWTPI  
PGTHREKLLVCHSFVPLSAPRARMVTSVCGVLTLMVIVHEAAVWRAGEPTKEH  
VDGVIT"  
complement(join(15723. .16044, 16879. .16957, 17029. .17134,  
20536. .20580, 21421. .21627, 22505. .22627))  
/gene="C45E1.4"  
/note="C45E1.4"  
/codon\_start=1  
/protein\_id="AAC24284.1"  
/db\_xref="GI:2804424"  
/translation="MSSLGTYFFLLVLLPVCATCYEEDYRPEGLTGHNGVFOALP  
WTKFELNLISSLHATANYPEVNLVRERMIISDIAPNDRRKIERMLKNLRPPVDEF  
LTEDETKVRKAHSERDVSILMNLNVLVIVVHSLKWEVKKIIGEHNFVDSN  
IRFTGLTDSKYIFLDLSLKHSHQSDLDSPSHCQCTDLKWEKFWSPSKELELQ  
KRKWCCTMEIIFEVKIYFSVKYKTFQVVCVQIKSFLSGKFDKFAAYPWTPI  
PGTHREKLLVCHSFVPLSAPRARMVTSVCGVLTLMVIVHEAAVWRAGEPTKEH  
VDGVIT"  
complement(join(15723. .16044, 16879. .16957, 17029. .17134,  
20536. .20580, 21421. .21627, 22505. .22627))  
/gene="C45E1.4"  
/note="C45E1.4"  
/codon\_start=1  
/protein\_id="AAC24284.1"  
/db\_xref="GI:2804424"  
/translation="MSSLGTYFFLLVLLPVCATCYEEDYRPEGLTGHNGVFOALP  
WTKFELNLISSLHATANYPEVNLVRERMIISDIAPNDRRKIERMLKNLRPPVDEF  
LTEDETKVRKAHSERDVSILMNLNVLVIVVHSLKWEVKKIIGEHNFVDSN  
IRFTGLTDSKYIFLDLSLKHSHQSDLDSPSHCQCTDLKWEKFWSPSKELELQ  
KRKWCCTMEIIFEVKIYFSVKYKTFQVVCVQIKSFLSGKFDKFAAYPWTPI  
PGTHREKLLVCHSFVPLSAPRARMVTSVCGVLTLMVIVHEAAVWRAGEPTKEH  
VDGVIT"  
complement(join(15723. .16044, 16879. .16957, 17029. .17134,  
20536. .20580, 21421. .21627, 22505. .22627))  
/gene="C45E1.4"  
/note="C45E1.4"  
/codon\_start=1  
/protein\_id="AAC24284.1"  
/db\_xref="GI:2804424"  
/translation="MSSLGTYFFLLVLLPVCATCYEEDYRPEGLTGHNGVFOALP  
WTKFELNLISSLHATANYPEVNLVRERMIISDIAPNDRRKIERMLKNLRPPVDEF  
LTEDETKVRKAHSERDVSILMNLNVLVIVVHSLKWEVKKIIGEHNFVDSN  
IRFTGLTDSKYIFLDLSLKHSHQSDLDSPSHCQCTDLKWEKFWSPSKELELQ  
KRKWCCTMEIIFEVKIYFSVKYKTFQVVCVQIKSFLSGKFDKFAAYPWTPI  
PGTHREKLLVCHSFVPLSAPRARMVTSVCGVLTLMVIVHEAAVWRAGEPTKEH  
VDGVIT"  
complement(join(15723. .16044, 16879. .16957, 17029. .17134,  
20536. .20580, 21421. .21627, 22505. .22627))  
/gene="C45E1.4"  
/note="C45E1.4"  
/codon\_start=1  
/protein\_id="AAC24284.1"  
/db\_xref="GI:2804424"  
/translation="MSSLGTYFFLLVLLPVCATCYEEDYRPEGLTGHNGVFOALP  
WTKFELNLISSLHATANYPEVNLVRERMIISDIAPNDRRKIERMLKNLRPPVDEF  
LTEDETKVRKAHSERDVSILMNLNVLVIVVHSLKWEVKKIIGEHNFVDSN  
IRFTGLTDSKYIFLDLSLKHSHQSDLDSPSHCQCTDLKWEKFWSPSKELELQ  
KRKWCCTMEIIFEVKIYFSVKYKTFQVVCVQIKSFLSGKFDKFAAYPWTPI  
PGTHREKLLVCHSFVPLSAPRARMVTSVCGVLTLMVIVHEAAVWRAGEPTKEH  
VDGVIT"  
complement(join(15723. .16044, 16879. .16957, 17029. .17134,  
20536. .20580, 21421. .21627, 22505. .22627))  
/gene="C45E1.4"  
/note="C45E1.4"  
/codon\_start=1  
/protein\_id="AAC24284.1"  
/db\_xref="GI:2804424"  
/translation="MSSLGTYFFLLVLLPVCATCYEEDYRPEGLTGHNGVFOALP  
WTKFELNLISSLHATANYPEVNLVRERMIISDIAPNDRRKIERMLKNLRPPVDEF  
LTEDETKVRKAHSERDVSILMNLNVLVIVVHSLKWEVKKIIGEHNFVDSN  
IRFTGLTDSKYIFLDLSLKHSHQSDLDSPSHCQCTDLKWEKFWSPSKELELQ  
KRKWCCTMEIIFEVKIYFSVKYKTFQVVCVQIKSFLSGKFDKFAAYPWTPI  
PGTHREKLLVCHSFVPLSAPRARMVTSVCGVLTLMVIVHEAAVWRAGEPTKEH  
VDGVIT"  
complement(join(15723. .16044, 16879. .16957, 17029. .17134,  
20536. .20580, 21421. .21627, 22505. .22627))  
/gene="C45E1.4"  
/note="C45E1.4"  
/codon\_start=1  
/protein\_id="AAC24284.1"  
/db\_xref="GI:2804424"  
/translation="MSSLGTYFFLLVLLPVCATCYEEDYRPEGLTGHNGVFOALP  
WTKFELNLISSLHATANYPEVNLVRERMIISDIAPNDRRKIERMLKNLRPPVDEF  
LTEDETKVRKAHSERDVSILMNLNVLVIVVHSLKWEVKKIIGEHNFVDSN  
IRFTGLTDSKYIFLDLSLKHSHQSDLDSPSHCQCTDLKWEKFWSPSKELELQ  
KRKWCCTMEIIFEVKIYFSVKYKTFQVVCVQIKSFLSGKFDKFAAYPWTPI  
PGTHREKLLVCHSFVPLSAPRARMVTSVCGVLTLMVIVHEAAVWRAGEPTKEH  
VDGVIT"  
complement(join(15723. .16044, 16879. .16957, 17029. .17134,  
20536. .20580, 21421. .21627, 22505. .22627))  
/gene="C45E1.4"  
/note="C45E1.4"  
/codon\_start=1  
/protein\_id="AAC24284.1"  
/db\_xref="GI:2804424"  
/translation="MSSLGTYFFLLVLLPVCATCYEEDYRPEGLTGHNGVFOALP  
WTKFELNLISSLHATANYPEVNLVRERMIISDIAPNDRRKIERMLKNLRPPVDEF  
LTEDETKVRKAHSERDVSILMNLNVLVIVVHSLKWEVKKIIGEHNFVDSN  
IRFTGLTDSKYIFLDLSLKHSHQSDLDSPSHCQCTDLKWEKFWSPSKELELQ  
KRKWCCTMEIIFEVKIYFSVKYKTFQVVCVQIKSFLSGKFDKFAAYPWTPI  
PGTHREKLLVCHSFVPLSAPRARMVTSVCGVLTLMVIVHEAAVWRAGEPTKEH  
VDGVIT"  
complement(join(15723. .16044, 16879. .16957, 17029. .17134,  
20536. .20580, 21421. .21627, 22505. .22627))  
/gene="C45E1.4"  
/note="C45E1.4"  
/codon\_start=1  
/protein\_id="AAC24284.1"  
/db\_xref="GI:2804424"  
/translation="MSSLGTYFFLLVLLPVCATCYEEDYRPEGLTGHNGVFOALP  
WTKFELNLISSLHATANYPEVNLVRERMIISDIAPNDRRKIERMLKNLRPPVDEF  
LTEDETKVRKAHSERDVSILMNLNVLVIVVHSLKWEVKKIIGEHNFVDSN  
IRFTGLTDSKYIFLDLSLKHSHQSDLDSPSHCQCTDLKWEKFWSPSKELELQ  
KRKWCCTMEIIFEVKIYFSVKYKTFQVVCVQIKSFLSGKFDKFAAYPWTPI  
PGTHREKLLVCHSFVPLSAPRARMVTSVCGVLTLMVIVHEAAVWRAGEPTKEH  
VDGVIT"  
complement(join(15723. .16044, 16879. .16957, 17029. .17134,  
20536. .20580, 21421. .21627, 22505. .22627))  
/gene="C45E1.4"  
/note="C45E1.4"  
/codon\_start=1  
/protein\_id="AAC24284.1"  
/db\_xref="GI:2804424"  
/translation="MSSLGTYFFLLVLLPVCATCYEEDYRPEGLTGHNGVFOALP  
WTKFELNLISSLHATANYPEVNLVRERMIISDIAPNDRRKIERMLKNLRPPVDEF  
LTEDETKVRKAHSERDVSILMNLNVLVIVVHSLKWEVKKIIGEHNFVDSN  
IRFTGLTDSKYIFLDLSLKHSHQSDLDSPSHCQCTDLKWEKFWSPSKELELQ  
KRKWCCTMEIIFEVKIYFSVKYKTFQVVCVQIKSFLSGKFDKFAAYPWTPI  
PGTHREKLLVCHSFVPLSAPRARMVTSVCGVLTLMVIVHEAAVWRAGEPTKEH  
VDGVIT"  
complement(join(15723. .16044, 16879. .16957, 17029. .17134,  
20536. .20580, 21421. .21627, 22505. .22627))  
/gene="C45E1.4"  
/note="C45E1.4"  
/codon\_start=1  
/protein\_id="AAC24284.1"  
/db\_xref="GI:2804424"  
/translation="MSSLGTYFFLLVLLPVCATCYEEDYRPEGLTGHNGVFOALP  
WTKFELNLISSLHATANYPEVNLVRERMIISDIAPNDRRKIERMLKNLRPPVDEF  
LTEDETKVRKAHSERDVSILMNLNVLVIVVHSLKWEVKKIIGEHNFVDSN  
IRFTGLTDSKYIFLDLSLKHSHQSDLDSPSHCQCTDLKWEKFWSPSKELELQ  
KRKWCCTMEIIFEVKIYFSVKYKTFQVVCVQIKSFLSGKFDKFAAYPWTPI  
PGTHREKLLVCHSFVPLSAPRARMVTSVCGVLTLMVIVHEAAVWRAGEPTKEH  
VDGVIT"  
complement(join(15723. .16044, 16879. .16957, 17029. .17134,  
20536. .20580, 21421. .21627, 22505. .22627))  
/gene="C45E1.4"  
/note="C45E1.4"  
/codon\_start=1  
/protein\_id="AAC24284.1"  
/db\_xref="GI:2804424"  
/translation="MSSLGTYFFLLVLLPVCATCYEEDYRPEGLTGHNGVFOALP  
WTKFELNLISSLHATANYPEVNLVRERMIISDIAPNDRRKIERMLKNLRPPVDEF  
LTEDETKVRKAHSERDVSILMNLNVLVIVVHSLKWEVKKIIGEHNFVDSN  
IRFTGLTDSKYIFLDLSLKHSHQSDLDSPSHCQCTDLKWEKFWSPSKELELQ  
KRKWCCTMEIIFEVKIYFSVKYKTFQVVCVQIKSFLSGKFDKFAAYPWTPI  
PGTHREKLLVCHSFVPLSAPRARMVTSVCGVLTLMVIVHEAAVWRAGEPTKEH  
VDGVIT"  
complement(join(15723. .16044, 16879. .16957, 17029. .17134,  
20536. .20580, 21421. .21627, 22505. .22627))  
/gene="C45E1.4"  
/note="C45E1.4"  
/codon\_start=1  
/protein\_id="AAC24284.1"  
/db\_xref="GI:2804424"  
/translation="MSSLGTYFFLLVLLPVCATCYEEDYRPEGLTGHNGVFOALP  
WTKFELNLISSLHATANYPEVNLVRERMIISDIAPNDRRKIERMLKNLRPPVDEF  
LTEDETKVRKAHSERDVSILMNLNVLVIVVHSLKWEVKKIIGEHNFVDSN  
IRFTGLTDSKYIFLDLSLKHSHQSDLDSPSHCQCTDLKWEKFWSPSKELELQ  
KRKWCCTMEIIFEVKIYFSVKYKTFQVVCVQIKSFLSGKFDKFAAYPWTPI  
PGTHREKLLVCHSFVPLSAPRARMVTSVCGVLTLMVIVHEAAVWRAGEPTKEH  
VDGVIT"  
complement(join(15723. .16044, 16879. .16957, 17029. .17134,  
20536. .20580, 21421. .21627, 22505. .22627))  
/gene="C45E1.4"  
/note="C45E1.4"  
/codon\_start=1  
/protein\_id="AAC24284.1"  
/db\_xref="GI:2804424"  
/translation="MSSLGTYFFLLVLLPVCATCYEEDYRPEGLTGHNGVFOALP  
WTKFELNLISSLHATANYPEVNLVRERMIISDIAPNDRRKIERMLKNLRPPVDEF  
LTEDETKVRKAHSERDVSILMNLNVLVIVVHSLKWEVKKIIGEHNFVDSN  
IRFTGLTDSKYIFLDLSLKHSHQSDLDSPSHCQCTDLKWEKFWSPSKELELQ  
KRKWCCTMEIIFEVKIYFSVKYKTFQVVCVQIKSFLSGKFDKFAAYPWTPI  
PGTHREKLLVCHSFVPLSAPRARMVTSVCGVLTLMVIVHEAAVWRAGEPTKEH  
VDGVIT"  
complement(join(15723. .16044, 16879. .16957, 17029. .17134,  
20536. .20580, 21421. .21627, 22505. .22627))  
/gene="C45E1.4"  
/note="C45E1.4"  
/codon\_start=1  
/protein\_id="AAC24284.1"  
/db\_xref="GI:2804424"  
/translation="MSSLGTYFFLLVLLPVCATCYEEDYRPEGLTGHNGVFOALP  
WTKFELNLISSLHATANYPEVNLVRERMIISDIAPNDRRKIERMLKNLRPPVDEF  
LTEDETKVRKAHSERDVSILMNLNVLVIVVHSLKWEVKKIIGEHNFVDSN  
IRFTGLTDSKYIFLDLSLKHSHQSDLDSPSHCQCTDLKWEKFWSPSKELELQ  
KRKWCCTMEIIFEVKIYFSVKYKTFQVVCVQIKSFLSGKFDKFAAYPWTPI  
PGTHREKLLVCHSFVPLSAPRARMVTSVCGVLTLMVIVHEAAVWRAGEPTKEH  
VDGVIT"  
complement(join(15723. .16044, 16879. .16957, 17029. .17134,  
20536. .20580, 21421. .21627, 22505. .22627))  
/gene="C45E1.4"  
/note="C45E1.4"  
/codon\_start=1  
/protein\_id="AAC24284.1"  
/db\_xref="GI:2804424"  
/translation="MSSLGTYFFLLVLLPVCATCYEEDYRPEGLTGHNGVFOALP  
WTKFELNLISSLHATANYPEVNLVRERMIISDIAPNDRRKIERMLKNLRPPVDEF  
LTEDETKVRKAHSERDVSILMNLNVLVIVVHSLKWEVKKIIGEHNFVDSN  
IRFTGLTDSKYIFLDLSLKHSHQSDLDSPSHCQCTDLKWEKFWSPSKELELQ  
KRKWCCTMEIIFEVKIYFSVKYKTFQVVCVQIKSFLSGKFDKFAAYPWTPI  
PGTHREKLLVCHSFVPLSAPRARMVTSVCGVLTLMVIVHEAAVWRAGEPTKEH  
VDGVIT"  
complement(join(15723. .16044, 16879. .16957, 17029. .17134,  
20536. .20580, 21421. .21627, 22505. .22627))  
/gene="C45E1.4"  
/note="C45E1.4"  
/codon\_start=1  
/protein\_id="AAC24284.1"  
/db\_xref="GI:2804424"  
/translation="MSSLGTYFFLLVLLPVCATCYEEDYRPEGLTGHNGVFOALP  
WTKFELNLISSLHATANYPEVNLVRERMIISDIAPNDRRKIERMLKNLRPPVDEF  
LTEDETKVRKAHSERDVSILMNLNVLVIVVHSLKWEVKKIIGEHNFVDSN  
IRFTGLTDSKYIFLDLSLKHSHQSDLDSPSHCQCTDLKWEKFWSPSKELELQ  
KRKWCCTMEIIFEVKIYFSVKYKTFQVVCVQIKSFLSGKFDKFAAYPWTPI  
PGTHREKLLVCHSFVPLSAPRARMVTSVCGVLTLMVIVHEAAVWRAGEPTKEH  
VDGVIT"  
complement(join(15723. .16044, 16879. .16957, 17029. .17134,  
20536. .20580, 21421. .21627, 22505. .22627))  
/gene="C45E1.4"  
/note="C45E1.4"  
/codon\_start=1  
/protein\_id="AAC24284.1"  
/db\_xref="GI:2804424"  
/translation="MSSLGTYFFLLVLLPVCATCYEEDYRPEGLTGHNGVFOALP  
WTKFELNLISSLHATANYPEVNLVRERMIISDIAPNDRRKIERMLKNLRPPVDEF  
LTEDETKVRKAHSERDVSILMNLNVLVIVVHSLKWEVKKIIGEHNFVDSN  
IRFTGLTDSKYIFLDLSLKHSHQSDLDSPSHCQCTDLKWEKFWSPSKELELQ  
KRKWCCTMEIIFEVKIYFSVKYKTFQVVCVQIKSFLSGKFDKFAAYPWTPI  
PGTHREKLLVCHSFVPLSAPRARMVTSVCGVLTLMVIVHEAAVWRAGEPTKEH  
VDGVIT"  
complement(join(15723. .16044, 16879. .16957, 17029. .17134,  
20536. .20580, 21421. .21627, 22505. .22627))  
/gene="C45E1.4"  
/note="C45E1.4"  
/codon\_start=1  
/protein\_id="AAC24284.1"  
/db\_xref="GI:2804424"  
/translation="MSSLGTYFFLLVLLPVCATCYEEDYRPEGLTGHNGVFOALP  
WTKFELNLISSLHATANYPEVNLVRERMIISDIAPNDRRKIERMLKNLRPPVDEF  
LTEDETKVRKAHSERDVSILMNLNVLVIVVHSLKWEVKKIIGEHNFVDSN  
IRFTGLTDSKYIFLDLSLKHSHQSDLDSPSHCQCTDLKWEKFWSPSKELELQ  
KRKWCCTMEIIFEVKIYFSVKYKTFQVVCVQIKSFLSGKFDKFAAYPWTPI  
PGTHREKLLVCHSFVPLSAPRARMVTSVCGVLTLMVIVHEAAVWRAGEPTKEH  
VDGVIT"  
complement(join(15723. .16044, 16879. .16957, 17029. .17134,  
20536. .20580, 21421. .21627, 22505. .22627))  
/gene="C45E1.4"  
/note="C45E1.4"  
/codon\_start=1  
/protein\_id="AAC24284.1"  
/db\_xref="GI:2804424"  
/translation="MSSLGTYFFLLVLLPVCATCYEEDYRPEGLTGHNGVFOALP  
WTKFELNLISSLHATANYPEVNLVRERMIISDIAPNDRRKIERMLKNLRPPVDEF  
LTEDETKVRKAHSERDVSILMNLNVLVIVVHSLKWEVKKIIGEHNFVDSN  
IRFTGLTDSKYIFLDLSLKHSHQSDLDSPSHCQCTDLKWEKFWSPSKELELQ  
KRKWCCTMEIIFEVKIYFSVKYKTFQVVCVQIKSFLSGKFDKFAAYPWTPI  
PGTHREKLLVCHSFVPLSAPRARMVTSVCGVLTLMVIVHEAAVWRAGEPTKEH  
VDGVIT"  
complement(join(15723. .16044, 16879. .16957, 17029. .17134,  
20536. .20580, 21421. .21627, 22505. .22627))  
/gene="C45E1.4"  
/note="C45E1.



```

SOURCE
ORGANISM
Caenorhabditis elegans.
Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditidae;
Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE
AUTHORS
Waterston, R.H.
TITLE
The sequence of Caenorhabditis elegans clone
JOURNAL
Unpublished
AUTHORS
2 (bases 1 to 33673)
Waterston, R.H.
TITLE
Direct Submission
JOURNAL
Submitted (23-FEB-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
COMMENT
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 33673: contig of 33673 bp in length.
*
FEATURES
Location/Qualifiers
source
1. .33673
/organism="Caenorhabditis elegans"
/db_xref="taxon:6239"
/clone="K08D12"
BASE COUNT 10444 a 6365 c 6792 g 10074 t
ORIGIN

alignment_scores:
Quality: 51.00 Length: 20
Ratio: 3.188 Gaps: 0
Percent Similarity: 80.000 Percent Identity: 55.000

alignment_block:
US-08-653-294-31 x AC006672 ..
Align seg 1/1 to: AC006672 from: 1 to: 33673

1 TyrArgLeuAlaIleArgLeuAsnGluArgTyrArgLeuAlaIleArgLe 17
3872 CATAGGCTTGGCTTAGGCTTAAGCTTAGAGTTAGGCTTAGGCT 3921
17 uAsnGluArg 20
1::: |||
3922 TCGGCCTAGG 3931

seq_name: gb_htg2:AC006842

seq_documentation_block:
LOCUS AC006842 299015 bp DNA HTG 24-FEB-1999
DEFINITION Caenorhabditis elegans clone Y104H12X, *** SEQUENCING IN PROGRESS
***, 13 unordered pieces.
ACCESSION AC006842
VERSION AC006842.1 GI:4263506
KEYWORDS HTG; HTGS_PHASE1.
SOURCE
ORGANISM
Caenorhabditis elegans.
Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditidae;
Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE
AUTHORS
Waterston, R.H.
TITLE
The sequence of Caenorhabditis elegans clone
JOURNAL
Unpublished
AUTHORS
2 (bases 1 to 299015)
Waterston, R.H.
TITLE
Direct Submission
JOURNAL
Submitted (24-FEB-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA

```

---

```

COMMENT
* NOTE: This is a 'working draft' sequence. It currently
* consists of 13 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 1993: contig of 1993 bp in length
* 2009: gap of unknown length
* 4578: contig of 2569 bp in length
* 4579: gap of unknown length
* 4595: contig of 2682 bp in length
* 7276: contig of 2682 bp in length
* 7277: gap of unknown length
* 7293: contig of 4213 bp in length
* 11505: contig of unknown length
* 11521: gap of unknown length
* 11506: contig of 5417 bp in length
* 11522: gap of unknown length
* 16939: contig of 3980 bp in length
* 16934: gap of unknown length
* 20935: gap of unknown length
* 20951: contig of 4235 bp in length
* 25185: contig of unknown length
* 25201: gap of unknown length
* 25186: contig of 23959 bp in length
* 25202: gap of unknown length
* 49161: contig of 15797 bp in length
* 49177: gap of unknown length
* 64974: contig of 36565 bp in length
* 64990: gap of unknown length
* 101555: contig of 56116 bp in length
* 101570: gap of unknown length
* 157686: contig of 102657 bp in length
* 157702: gap of unknown length
* 157703: contig of 102657 bp in length
* 260360: gap of unknown length
* 260376: contig of 38640 bp in length.
*
FEATURES
Location/Qualifiers
source
1. .299015
/organism="Caenorhabditis elegans"
/db_xref="taxon:6239"
/clone="Y104H12X"
BASE COUNT 96423 a 53226 c 53209 g 95965 t 192 others
ORIGIN

alignment_scores:
Quality: 51.00 Length: 20
Ratio: 3.188 Gaps: 0
Percent Similarity: 80.000 Percent Identity: 55.000

alignment_block:
US-08-653-294-31 x AC006842 ..
Align seg 1/1 to: AC006842 from: 1 to: 299015

1 TyrArgLeuAlaIleArgLeuAsnGluArgTyrArgLeuAlaIleArgLe 17
69270 CATAGGCTTGGCTTAGGCTTAAGCTTAGAGTTAGGCTTAGGCTTAGGCT 69319
17 uAsnGluArg 20
1::: |||
69320 TCGGCCTAGG 69329

seq_name: gb_htg2:AC006874

seq_documentation_block:
LOCUS AC006874 306131 bp DNA HTG 24-FEB-1999
DEFINITION Caenorhabditis elegans clone Y54E10, *** SEQUENCING IN PROGRESS
***, 3 unordered pieces.
ACCESSION AC006874
VERSION AC006874.1 GI:4263474
KEYWORDS HTG; HTGS_PHASE1.
SOURCE
ORGANISM
Caenorhabditis elegans.
Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditidae;
Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.

```



11454 TAGGCTTAGG 11445

**THIS PAGE BLANK (USPTO)**